Foreign Affairs and
International Trade Canada

Affaires étrangères et
Commerce international Canada

Government
Publications

Canada

Trade Policy
Research
2005

Trade Policy
Research
2005

John M. Curtis and Dan Ciuriak
Editors

Table of Contents

Foreword

This volume brings together the results of analysis and policy research undertaken within, on behalf of, or in collaboration with Foreign Affairs and International Trade Canada over the past year. Launched in 2001 as part of the response to the Government of Canada's *Policy Research Initiative*, a government-wide effort to re-create and expand its research capacity, the *Trade Policy Research* series is now in its fifth edition.

Previous volumes have traced the debate in trade policy circles since the watershed developments at the 1999 WTO Ministerial in Seattle, following the launch and hesitant progress of the Doha Round, touching on topical issues and showcasing research and analysis conducted within the Government of Canada on various aspects of trade policy and economic globalization more generally.

This year's volume continues in that vein. Part I addresses issues confronting the Doha Development Agenda as it moves towards the closing phase of negotiations.

Part II compiles several staff papers and commentaries completed in 2005. The seven chapters describe: the determinants of the direction of Canada's direct investment flows abroad; the dynamic version of Foreign Affairs and International Trade Canada's computable general equilibrium model for analyzing the impact of trade policies on Canada; an adaptation of a standard partial equilibrium model for use within the Department; a method of directly including revealed comparative advantage in a gravity model of trade; a preliminary assessment of the impact of embassies and consulates on Canadian trade performance; and two commentaries on topical aspects of trade policy.

Through this volume, Foreign Affairs and International Trade Canada seeks to continue to contribute actively to the discussion concerning the role of international trade and investment in Canada and in the global economy more generally. And, in the process, we continue to work in the spirit of the

broader commitment of the Government of Canada to stimulate the development of its applied research capacity. At the same time, we continue to foster links with professional and academic commentators by continuing the pattern set in previous *Trade Policy Research* editions of including contributions from that sector. Accordingly, the papers presented in the various chapters are written in the personal capacity of the authors and do not represent the views of the Department, of the Government of Canada, or of any organizations with which the respective authors might be affiliated.

This volume was produced under the guidance of John M. Curtis, Chief Economist, Foreign Affairs and International Trade Canada, together with co-editor and managing editor Dan Ciuriak, Deputy Chief Economist, Foreign Affairs and International Trade Canada.

Ottawa
March, 2006

Part I

Multilateral Trade Issues

Scoping the End-Game in the Doha Round: A Roundtable Discussion

John M. Curtis and Dan Ciuriak[*]

On February 17-18, 2005, the Centre for International Governance Innovation at the University of Waterloo and the International Development Research Centre, in co-ordination with the Department of Foreign Affairs and International Trade convened an informal meeting of leading observers of the international trade and investment scene for a discussion of the progress of the Doha Development Agenda as the focus shifted to the scoping out of the "end game" for the negotiations and more particularly to the preparations for the Sixth Ministerial Conference of the World Trade Organization (WTO) at Hong Kong, China, December 2005. The goal of the roundtable discussion was to share views on the progress of negotiations and the prospects for an ambitious outcome, taking into account the geopolitical and international macroeconomic context; to discuss issues that might usefully be addressed given the directions being taken in the negotiations; and to identify areas where analytic work might facilitate further progress. This note represents the Chair's thematic summary of the discussions. As these were held under Chatham House rules, no attribution is given. Responsibility for the interpretation of the discussion rests entirely with the authors.

1

Introduction

Now in its fourth year, the Doha Development Agenda, the ninth round of multilateral trade negotiations under the auspices of the GATT/WTO, is already well on its way to being as storied as its immediate predecessor, the Uruguay Round, which took eight years and the overcoming of many trials and tribulations to complete. After a surprising plot twist which saw the Doha Round revived from an ignominious collapse at Cancún in September 2003 by an agreement on negotiating modalities in Geneva in July 2004, the questions of the moment have become, first, whether negotiators can take advantage of renewed life in the Round to bring it to completion; and second, whether an outcome in the present Round can be built up to match that of its predecessor.

In good measure, the rather pedestrian answer to these questions is that much will be determined by the extent of progress in the negotiations during 2005 and, most importantly, by the results of the December 2005 Hong Kong ministerial meeting. A successful outcome in this context will depend on factors as varied as:

- The preparation of a good first approximation of the text of a potential final agreement by mid-2005.
- The avoidance of fractiousness in the selection of a new Director General of the WTO to replace the incumbent, Supachai Panitchpakdi, whose term expires on 1 September 2005.[1]
- A smooth assumption of the burdens of trade leadership by Robert Portman and Peter Mandelson, the successors of US Trade Representative Robert Zoellick and EU Trade Commissioner Pascal Lamy, two individuals who combined tireless advocacy of trade liberalization with the technician's command of the substance of negotiations.
- Substantively, finding a way to unlock meaningful offers on agriculture, services and non-agricultural manufactured

---

[1] The transition was indeed smooth: Pascal Lamy became the fifth Director-General of the WTO, as of 1 September 2005 for a four-year term.

goods in order to move the negotiations from discussion of modalities to discussion of commercially meaningful market opening—for developed and developing countries alike.

While good progress in 2005 would pave the way for the final run to an agreement, it would not of course be the end of the story. A number of other shoes would still have to fall in the end-game before the Doha Round would be safely in the books.

- A credible commitment of resources for trade-related technical assistance and capacity building would have to be forthcoming to keep the large bloc of poor countries (the G-90[2]) at the table.
- Solutions would have to be at hand for the systemic and institutional issues that are likely to arise, as they typically have in the past, in the late stages of the negotiations to help balance the deal.
- The various protagonists in the negotiating process would have to meet the procedural challenges of bringing along each of the negotiating groups, of aggregating the results in understandable ways for the Members to facilitate the calculation of the trade-offs that they individually face, and of maintaining transparency in the process to forestall an inadvertent breakdown of trust that could derail the process altogether.

Are there innovative ideas on ways to break through the various impasses that have slowed or stalled negotiations to date? Will the major industrialized states see multilateral trade liberalization as a solution to their current macroeconomic problems—or as an impediment or even as irrelevant? Will the large dynamic developing countries that are presently doing well in economic terms complacently decide to live with the *status quo*, or make essential contributions to the trading system from which they have drawn much of their recent dynamism? Will international business enterprises see improvements to the operation of the multilateral system as important to their bottom lines?

---

[2] The G-90 is a WTO-oriented coalition of the least developed countries (LDCs), the African Union (AU) and the African, Caribbean and Pacific (ACP) group.

As we will see from the discussion recounted below, the prospects are not without hope, but there are no clear answers to these questions, which means that 2005 will, for those involved in trade policy, be a year of living rather nervously.

The macroeconomic and geopolitical context

Historically, many have argued, the macroeconomic context has played an important role in the progress of trade rounds as policymakers looked to trade to help resolve imbalances.[3] This time around, it was pointed out by some observers, the linkage is not being made. Not even by markets.

Global imbalances are not driving trade activism...

In the shadow of global macroeconomic imbalances the sheer size of which is straining the imagination of the economic commentariat if not yet overtaxing the nerves of international investors, 2004 was the best year for global growth in several years. The macroeconomic imbalances have to date been benign and, going forward, are not expected by the mainstream to lead to precipitous adjustments.

As regards exchange rates, some 50 percent of the adjustment in exchange rate parities thought by some to be needed to help turn the dynamic towards reducing imbalances had already happened by the end of 2004, albeit in an unbalanced way, with the US dollar depreciating against the euro and the floating commodity currencies but with limited participation from Asia. The key to the remaining part of the dollar adjustment was thought to be China moving to a more flexible exchange rate regime, which it had long signaled it would do when the re-

---

[3] See, for example, C. Fred Bergsten, "The International Monetary Scene and the Next WTO Negotiations", in Jeffrey J. Schott, *Launching New Global Trade Talks: An Action Agenda*, (Washington: Institute for International Economics, September 1998): 39-46. Also see John M. Curtis. "The Role of Contextual Factors in the Launching of Trade Rounds", in John M. Curtis and Dan Ciuriak (eds.) *Trade Policy Research 2002* (Ottawa: Department of Foreign Affairs and International Trade, 2002): 33-70.

forms to its banking system were sufficiently advanced, the timing of which remained uncertain.[4]

As regards the US domestic savings-investment imbalance which has been exacerbated by the expansion of the US fiscal deficit and which has increasingly come to be seen as to blame for the external imbalance, it was noted that there is a general expectation of some improvement with robust global economic growth reducing the cyclical component in the near term and the Administration having announced a program of fiscal consolidation to reduce the structural component in the medium term.

However, it was suggested, the so-called "twin deficit" issue has not generated a sense of urgency in Washington.[5] In part, this reflects consideration of the context in which the US current account deficit originated: some two-thirds of the external deficit was built up during the late 1990s at a time when the US was going into fiscal surplus. This was a period of high growth in US productivity and of the US economy overall, while slow growth in continental Europe and Japan weakened external demand for US exports. This suggested that cyclical factors were primarily responsible. And since the expansion of the external and fiscal deficits has *not* led to rising longer-term interest rates, typically a symptom of investor unease, the more compelling case to some analysts is that the external deficit reflects the attractiveness of the United States to foreign capital—in other words, the US capital account surplus is driving the current account deficit. Moreover, insofar as there *is* a real fiscal problem in the United States, it was thought to lie in the longer term as

[4] Editors' note: On 21 July, 2005, China announced that its currency, the renmimbi (RMB), would no longer be pegged to the US dollar, but would float against a basket of currencies. The initial revaluation of the RMB by 2.1 percent from 8.28 to 8.11 was not however followed by further significant changes in its parity; the RMB traded at 8.09 to the US dollar in mid-October.

[5] In fiscal year 2004, the US federal government deficit, inclusive of social security, amounted to US$412.1 billion or 3.5 percent of 2004 US GDP of US$11,733.5 billion; the current account deficit in calendar year 2004 amounted to US$665.9 billion or 5.7 percent of GDP. Source: Congressional Budget Office and US Bureau of Economic Analysis.

the burden of repayment of the rising debt shifts to future generations. But those longer-term problems will not come home to roost for the current Administration or even (not unimportantly for Administration political strategists) for the next Administration. There was, thus, in the view of some observers, a strong sense presently of "free lunches" in Washington.

The real test will come, it was argued, if interest rates get high enough to cause pain. But, it was pointed out by others, interest rates have not become a problem even with questions being raised about the Administration's seriousness on budget consolidation after it added US$80 billion to defence spending a week after releasing the 2005-2006 budget. And, as long as that remains the case, there will be no macroeconomic problem deriving from the budget deficit, or by extension from the external deficit.

...but might global imbalances pose risks to trade?

Even if one takes a benign view of the macroeconomic consequences of global imbalances (which many observers worldwide do not[6]), there are potential risks for the trade system from the external deficit. If there is no change in policy, current trends imply a continued widening of deficits.[7]

---

[6] As one observer put it, there are no precedents for the current situation; a gut-wrenching adjustment cannot be excluded. A detailed exposition of this view is provided by Lawrence H. Summers, "The U.S. Current Account Deficit and the Global Economy", *The Per Jacobsson Lecture*, October 3, 2004; http://www.perjacobsson.org/2004/100304.pdf.

[7] Conventional private sector macroeconomic projections have tended to show the US current account deficit flattening out in absolute terms and gradually declining as a share of GDP. For example, Global Insight projected the deficit growing from US$666 billion in 2004 to the US$700 billion range in 2005 and remaining essentially flat in absolute terms thereafter, while declining as a share of GDP from the 5.7 percent range in 2004 to the 5 percent range in 2007-2008. See Global Insight, *Quarterly Review and Outlook*, Fourth Quarter 2004. The continued widening of the deficit has substantially exceeded these expectations; in the first half of 2005, the deficit was close to US$800 billion at annual rates.

From the perspective of those who view exchange rate adjustment as the key, it was stressed that China holds that key—and China is doing quite well, thank you, and likes the current situation. And the rest of Asia likes China doing well. Moreover, whereas in the Uruguay Round, the active developing countries held out until the last minute but ultimately signed on to what many now regard as on balance a bad deal for them (including TRIPs, etc.), now China and others in the G-20 are prospering and do not have to sign on to any given deal. Even the Europeans, it was suggested, are mostly happy with things as they are. This implies continuation of the *status quo* (i.e., no rush to a deal and/or a small outcome at Hong Kong, China).

To be sure, this would generate ugly rhetoric in the United States—substitute China for Japan, suggested some, and the 1980s serves as a guide to the likely flavour of discussion in the United States over the coming months and likely the decade as a whole, although the actual pressure on China, the Schumer bill notwithstanding[8], cannot as yet be compared to the pressure on Japan in the 1980s.

But several argued that things have changed since the 1980s, especially in the economic and trade domains. To be sure, US power has not disappeared and, despite a long list of potential contenders, international relations are *not* operating on the basis of a balance of power. However, the WTO Agree-

---

[8] "S. 295: A bill to authorize appropriate action in the negotiations with the People's Republic of China regarding China's undervalued currency are not successful." The Bill was introduced February 3, 2005 by Senator Charles Schumer [D-NY], read twice and referred to the Committee on Finance. The Bill proposes a 27.5 percent tariff on all goods of Chinese provenance, that being the average of estimates considered by the Bill's framers of the undervaluation of China's currency, the renmimbi (RMB), against the US dollar. The Bill argues that the tariff would be consistent with the US' WTO commitments, basing it on Article XXI of the GATT which provides for measures taken for national security interest: "Protecting the United States manufacturing sector is essential to the interests of the United States." The tariff would remain in place until such time as China effected a substantial revaluation to bring the RMB into line with its "fair market value" and ceased to acquire foreign exchange reserves to prevent the appreciation of the RMB against the dollar.

ments have introduced a different legal regime to govern international trade while globalization has changed the business environment, including through the massive expansion of foreign direct investment and the organization of production in global value chains. Trade is increasingly accounted for by South-South flows; and the growing share of global growth accounted for by China means that at least some commodity prices are driven more by developments in China than in the United States. The South might thus be at least partially insulated from tensions in, or with, the North.

The geopolitical and geoeconomic shift in the power dynamic brought about by the expanded role in trade of Brazil, China and India has made itself felt in the WTO context. The traditional forum for trade leadership, the Quad, was described as effectively "finished". The ability of the US-EU partnership to drive the process has also been considerably diminished. For example, prior to Cancún, the impasse on agriculture led to a request by WTO Members to the United States and the European Union to develop a draft agreement; they agreed but Brazil said no, effectively blocking the agreement. Meanwhile, in the Geneva process leading up to the July 2004 agreement on a negotiating framework, the group that drove the process on agriculture was the Five Interested Parties (FIPs), which included Brazil and India alongside the United States and the European Union, with Australia as Cairns Group representative, and with Kenya being brought into the discussions to provide an African perspective. Pointedly, Chinese officials have privately described FIPs as "not large enough".

In the view of some, these developments underline the fact that, in the economic sphere, the United States now needs the cooperation of other nations to get things done. It cannot pressure them relentlessly on economic policy and is in any case hampered by the inward focus of its politics.

Some observers questioned just how important a priority trade would be for the present US Administration. Trade, it was noted, was not mentioned in the 2005 State of the Union address. While the United States was very active on trade during Robert Zoellick's tenure as United States Trade Representative

(USTR)—indeed it was argued that there would not be a Doha Round if Zoellick had not made it happen in the first place and then kept it alive in the Geneva process in 2004—there appears to be considerable risk that US trade leaders will not get backing from the domestic side when they push forward. The bilateral agreements that the US is negotiating, which were seen by some as merely a sideshow, were facing a tough time in Congress and were at risk of being shut down.[9] The inward focus in the United States was, as one observer noted, a rather ironic state of events given the global security situation.

While the path to resolution of the imbalances was as shrouded in uncertainty to roundtable participants as it appears to be to economic observers more generally, many tended to the not illogical perspective that unsustainable imbalances will not be sustained—and the only questions are what will trigger the correction and what will be the consequences?

There are enough possible triggers for an adjustment—events in the political sphere or in currency or financial markets—and the adjustment could start outside the United States: rising interest rates could be a problem in the developing countries, impacting negatively on sovereign debt, commodity prices and growth.

---

[9] Editors' note: The Dominican Republic-Central American free trade initiative (DR-CAFTA), which involved the United States, Costa Rica, Dominican Republic, El Salvador, Guatemala, Honduras and Nicaragua, faced intense lobbying against liberalizing sugar trade and was also slowed down by various other sticking points, the latest of which was an intellectual property issue with Guatemala which passed legislation to resolve the issue in March 2005. The DR-CAFTA eventually was passed by the US Congress, but by the narrowest of margins. The United States is also negotiating a separate pact with Panama and an FTA with the Andean countries (Colombia, Ecuador and Peru) in addition to the stalled Free Trade Area of the Americas (FTAA). Other US regional/bilateral initiatives include negotiating a number of FTAs in the Middle East that might be rolled up into a Middle East Free Trade Agreement (MEFTA), a proposed network of bilateral FTAs with ASEAN countries, and FTA talks with the South African Customs Union (SACU). A Congressional push is also mounting for an FTA with New Zealand.

The big complication, it was suggested, is the risk of a China-US power struggle emerging as the balance of power shifts and as the role of the US dollar comes under potential pressure.[10] In this regard, a comparison was drawn with Britain looking at the rise of German industrial power in the 1900s. Such tensions could spillover onto the trade system. For the smaller players in the international system, the talk of confrontation between the US and a rising China is unsettling.

Thus, much as it might be wished for by the trade policy community, a simplistic dynamic of macroeconomic policy problems leading to positive, forward-looking trade policy activism simply does not exist. Rather, there is the uncomfortable underlying sense that the combined macroeconomic/geopolitical context embodies risks for the trading system.

Still, it was argued, while the Doha Development Agenda may in the end be a minimal, face-saving round, the trading system will carry on business as usual. Will it?

Attenuating the risks to the system

Some observers sense a certain complacency about the stability of the global trade system with negotiations focused on the "nuts and bolts" while the major pressures that bear on the system go largely ignored. This is hardly unusual: the minutiae debated in

---

[10] A commentary noted in the roundtable discussions on the implications for the US of the shift in global power is that by Michael Lind, "How the U.S. Became the World's Dispensable Nation", *Financial Times*, 25 January 2005. Lind argues that "It is true that the US remains the only country capable of projecting military power throughout the world. But unipolarity in the military sphere, narrowly defined, is not preventing the rapid development of multipolarity in the geopolitical and economic arenas." In the latter regards, he suggests that "Today the evidence of foreign co-operation to reduce American primacy is everywhere – from the increasing importance of regional trade blocs that exclude the US to international space projects and military exercises in which the US is conspicuous by its absence." As regards foreign central banks diversifying away from the US dollar as reserve currency, there is some evidence that some have done this; however, such diversification has necessarily been made up for by the Chinese and Japanese central banks which have intervened heavily to support the dollar.

GATT circles as the Members were trying to launch the Uruguay Round reminded one observer of the last issue debated in the League of Nations just before World War II: level railway crossings. Is the WTO similarly at risk of losing itself in its own minutiae and missing the big picture?

Some argued, yes. Success is needed in the Doha Round negotiations for the good of the system—and a small round will not do. The Doha Round, it was suggested, better be on people's geopolitical agenda for several reasons:

- The pressures within the trade system are growing, not diminishing, and require the palliative offered by new negotiations. Not least of the sources of pressure are: (a) the expiry at the end of 2003 of the "peace clause" on agricultural disputes; (b) the use of the WTO dispute settlement understanding (DSU) for strategic purposes in order to set up subsequent negotiations; and (c) the negative sentiment in the US Congress to international courts.
- Regional/bilateral agreements are proliferating discriminatory measures.
- Trade was to provide the resources to help achieve the United Nations' Millennium Development goals. If the Doha Development Agenda fails, it might be many years before a new round could be put together. In the interim, it would not be possible to move forward on agriculture and services liberalization, both of which depend on the cross-sectoral trade-offs afforded by a comprehensive round. The broader development agenda was thus at risk.

Balancing the judicial and legislative roles of the WTO

Since the DSU is mandated to clarify the WTO Agreements, dispute settlement and negotiations are intended to work symbiotically within the WTO. The linkage between the WTO's judicial and legislative processes is emphasized by the occasional use of the DSU for strategic purposes in order to set up subsequent negotiations. Thus, in the absence of movement on the negotiating front, dispute settlement is subject to added pressure.

While no new complaints concerning agriculture have been raised since the expiry of the "peace clause", the Cairns Group could at any time, it was noted, immobilize the dispute settlement system with claims against agricultural subsidies under the WTO's subsidies code.[11] What strains would that raise, it was asked, for the trading system? The panel decision on cotton attacked so many US laws that, in the perspective of one observer, it was problematic from a systemic perspective: what if the United States were to lose several more such cases on agriculture or on the Farm Bill under the DSU? What would happen if the Appellate Body were to rule on the European Union's Common Agricultural Policy?

The dispute settlement system, including the research capacity in the WTO Secretariat, is already strained, many argued, by the high case load, the volume of which has vastly exceeded the expectations of the framers of the system.

To be sure, the cases taken to the panel stage are but a small part of the picture. Most international commercial disputes are settled outside the DSU altogether (e.g., through the food safety system, etc.). Moreover, diplomatic settlement early in the WTO dispute resolution system is quite active.[12] Former WTO Director General Mike Moore also emphasized the availability of provisions to use the good offices of the DG in mediation and conciliation, but with few takers to date. The Sutherland Report[13] sug-

---

[11] Editors' note: the two recent agriculture cases targeting sugar subsidies in the European Communities and US cotton subsidies were initiated before the expiry of the so called 'peace clause' under which WTO Members agreed not challenge domestic agricultural subsidies; the expiry of the peace clause was not a factor in either case, although it was referred to in the WTO hearings. No new agricultural cases have been brought to the WTO since the expiry of the 'peace clause'. See Bridges Weekly Trade News Digest, "DSB UPDATE: EC-SUGAR, US-COTTON" (Vol. 8 No. 10, 18 March 2004.)

[12] See, for example, Marc L. Busch and Eric Reinhardt, "The Evolution of GATT/WTO Dispute Settlement", in John M. Curtis and Dan Ciuriak (eds.) Trade Policy Research 2003 (Ottawa: Department of Foreign Affairs and International Trade, 2003): 143-183.

[13] Peter Sutherland, et al., "The Future of the WTO: Addressing institutional challenges in the new millennium", Report by the Consultative Board

gested the need to start seriously thinking about enlarging the role of the Secretariat and of the Director General.

Moreover, compliance is an issue. Overall, compliance has been very good with only some 8-10 cases where there has been an outright failure to comply or where compliance measures are subject to ongoing dispute. If there is a problem here, it is not with the small countries, which are complying. The potential problem lies with non-compliance by the big countries in big cases; such non-implementation, it was argued, is corrosive for the system.

Containing the risk element in regionalism

While all the major trade negotiations have slowed to a crawl or stalled, including the Doha Development Agenda, the Free Trade Area of the Americas (FTAA), and Mercosur/EU, smaller bilateral/regional agreements are proliferating, which in the view of some will inevitably weaken the multilateral system and the WTO—indeed, the "most favoured nation" (MFN) principle is in danger of being replaced by the "least favoured nation" or LFN principle as the volume of trade conducted under MFN shrinks. Every time the Doha Round appears to stall, the regional/bilateral dynamic intensifies.

IMF/World Bank analysis indicates that multilateral liberalization is the ideal route. While preferential trade agreements are generally considered to be "building blocks" towards freer trade, some argue that they do not result in a good environment for business. For example, the National Foreign Trade Council, a US business lobby, found that rules of origin (ROOs) arising from the plethora of preferential trade agreements were bad for its members. Accordingly, it was pushing for elimination of all duties on manufactured goods by 2015.[14]

to the Director-General Supachai Panitchpakdi (Switzerland: World Trade Organization, 2004).

[14] See Matthew J. Slaughter, "Tariff Elimination for Industrial Goods: Why the Gains Will Far Outweigh Any Losses", Background Paper prepared for the National Foreign Trade Council, August 2003, http://www.wto.org/english/forums_e/ngo_e/posp44_nftc_tarif_paper_e.pdf.

Moreover, preferences have proved to be of uncertain benefit even for those receiving them. The utilization of the General System of Preferences (GSP) is weak because of the operation of ROOs. For example, many apparel export contracts stipulate the source of textiles to be used in the manufacture of articles of clothing. Thus, if competitive East Asian textiles are used in South Asian apparel manufacture, the products cease to qualify under the GSP because of the GSP's ROOs.

Supporting development

Success is needed in the current multilateral trade talks, it was suggested, for broader reasons as well: in particular to support development. The Millennium Development Goals, many thought, will not be met ("not even close"). Trade was to have provided the resources and for this a minimal negotiated outcome simply will not do: subsidies must be cut in the rich countries, barriers must come down in middle income countries, reforms are needed in the poor countries. A non-ambitious outcome would not only be deleterious to the reform momentum in the poorer developing countries, it would also mean that aid would have to be maintained at higher levels than otherwise would be necessary. Accordingly, the aim of the Doha Round has to be pitched unambiguously high.

The path to an ambitious outcome

The question of the size of the outcome is not unrelated to the question of the path to the outcome. The challenge for 2005, in the estimation of some observers, was to set the level of ambition, which is implicit in the modalities. Every capital will have to decide what to put on the table. If countries bite the bullet on tough issues, it builds up the size of the outcome. The Round has been kept alive until now by marvelous work by key individuals such as Robert Zoellick, and movement on some key issues (in particular, the EU's pledge to eliminate agricultural export subsidies), but a quantum leap is needed to set up a successful conclusion to the Round.

Building the market access component

In an ideal world, offers are put on the table simultaneously. But if everyone waits to see what others will offer, we will have nothing. Who will take the lead? Not developing countries, it was argued; they have been burnt repeatedly by the failure of the industrialized countries to deliver on concessions and are now shy. This dates back to the Kennedy Round and was repeated in the Uruguay Round negotiations where they were given concessions yet saw that protection was put back up again through other instruments. China, Brazil and India are doing fine as things stand, which reduces the pressures on them to cut a deal. This leads some observers to conclude that the industrialized countries will again have to be the first movers.

Agriculture is the area where industrialized countries have a large margin of flexibility to provide the basis for a reasonable deal: some tariffs are sky high (e.g., one industrialized economy was cited as having some tariffs in the 700 percent range) and tariff rate quotas (TRQs) are often infinitesimal when considered as a share of the domestic market. The OECD's producer support estimate (PSE), a measure of public support for agriculture, shows that PSE levels have declined only from 36 percent in the mid-1980s to 31 percent now. Current public financial support for agriculture in the OECD countries, it was suggested on this basis, is larger than the GDP of Africa! Success in the Doha Development Agenda depends on very ambitious cuts to this level of support. In practical terms, this means that the US and the EU have to signal that they will provide something along the following lines:

(a) actual cuts in subsidies, not just in bound commitments (some estimates suggest that cuts have to amount to as much as 80 percent in order to provide meaningful liberalization); and

(b) improved market access in areas of particular export interest to developing economies.

In the opinion of some observers, we are close to the point of having strong signals as to the shape of the package on agriculture. There would be perhaps 70-80 percent cuts in bound

subsidies, although there remains some uncertainty about which products, and improved market access under TRQs. The timing of cuts to agriculture subsidies remains to be firmed up. But, if the United States were to offer something on counter-cyclical subsidies and the EU on export subsidies, that would constitute, in the estimation of some, the package.

Since agriculture cannot be self-balancing, the balance must come from services and non-agricultural market access (NAMA). A good offer on agriculture would, in the view of some, unlock the services and NAMA negotiations. Even without a rules component, a good market access package would constitute an overall successful outcome for the Round.

However, there has to be a market access *quid pro quo* for the industrialized countries. Although budget pressures might force a domestic consensus on subsidies to do what has to be done anyway, given the mercantilist ethos of trade negotiations, these countries need gains in services, areas of particular interest to multinational corporations, and something for their agricultural sectors—the "academic" approach to making the case for trade liberalization ("its good for you") does not, it was pointed out, get a negotiator to second base on Capitol Hill. It was also noted that even in the developing countries where much liberalization has been autonomous and driven by recognition that this was in the self-interest of the liberalizing country, the process has tended to stall when the pressure points are met. Reform must then be driven by outside pressures. Thus, India, Brazil and other emerging market economies need a successful round to drive further down the reform path that has led to their present success. This pushes the Round towards a bigger deal—which is essential to establish the very basis for a deal and moreover is needed by the WTO for the sake of its own sustainability, not least to stop the erosion of the MFN clause through proliferation of regional/bilateral agreements.

Accordingly, middle income developing countries have to offer meaningful market opening to the United States and to the poorer developing countries—although some would settle for developing countries simply binding the autonomous liberalization that they have already made in response to a sufficiently

attractive offer from the OECD countries on agriculture: as was pointed out, there have been reversals on liberalization, so binding autonomous liberalization would be a step forward (at the same time, the demandeurs on agricultural trade have yet to see a sufficiently attractive offer). Without those carrots of additional market access, domestic vested interests will prevent a deal. At the same time, pushing forward quickly on the Doha Round negotiations could inject a positive element into shaping the next US Farm Bill (the present one expires in 2007), which will be in the important preparatory stages in 2005-2006.

Not all agreed that agriculture is the key to unlocking the way forward on non-agricultural market access and services. While undoubtedly important for particular WTO Members, from the perspective both of an overall trade deal and of budget pressures in the major industrialized economies, the agricultural subsidies represent a comparatively minor part of the picture—in other words, they are neither *the* bargaining chip nor *the* source of pressure for autonomous change that some have made them out to be. Moreover, there is concern about the political will in the industrialized economies to move on agriculture: US action in the WTO Agriculture Committee's Sub-Committee on Cotton, while not unexpected, was unfortunately also not helpful from the perspective of the developing countries[15] and the signals on agriculture from ongoing US bilateral/regional negotiations were not seen by observers as encouraging (although, as was argued, this might largely reflect the modality: it is easier to negotiate incremental liberalization on agriculture in the multilateral context than to agree to free trade in a regional or bilateral agreement and then argue about the time frame to get there).

---

[15] Editors' note: At the 16 February 2005 meeting of the cotton sub-committee, the US opposed African and other least-developed WTO Members' position that the sub-committee's work program include a clear mandate for negotiations. In a subsequent proposal issued 25 February the developing countries proposed, *inter alia*, establishing modalities for market access, domestic support and export subsidies for the cotton negotiations. See " Differences Over Scope Of Work Programme Linger In WTO Cotton Sub-Committee", Bridges, *Weekly Trade News Digest*, Vol. 9, No. 7, 2 March 2005; http://www.ictsd.org/weekly/05-03-02/WTOinbrief.htm.

There was also skepticism about agriculture as the key to development aspects of the Round: accelerated reform of agriculture in India or Africa, it was suggested by one observer, would create endless numbers of landless peasants.

Similarly, there was skepticism about services liberalization. The services negotiations were considered to have gone nowhere; the services offers on the table were described as "worthless". As one observer put it, if *trade in services* were a stock, the recommendation would be to "buy"; if *negotiations* on trade in services were a stock, the recommendation would be to "sell".[16] The problem is intrinsic to the nature of services trade. Services offers are hard to make because there are implied domestic regulatory reforms that are hard to think through, let alone to act on due to domestic political considerations. It is simply unrealistic to expect developing countries to enter into binding agreements which could result in payments under dispute settlement when even the rich countries cannot fathom the true extent of their commitments. For example, a recent WTO panel decision found that the United States had in fact made GATS commitments in respect of gambling services that it claimed it had not intended.[17] While, upon appeal, the Appellate

---

[16] Editors' note: Corroborating this sense of the state of services negotiations, a senior WTO official reportedly described services as the "crisis item" on the agenda of the second WTO mini-ministerial meeting of 2005 in Mombasa, Kenya, at the beginning of March. See "Services, NAMA, Development Priorities At Kenya Mini-Ministerial", Bridges, *Weekly Trade News Digest*, Vol. 9, No. 7, 2 March 2005; http://www.ictsd.org/weekly/05-03-02/WTOinbrief.htm.

[17] *United States — Measures Affecting the Cross-Border Supply of Gambling and Betting Services: Report of the Panel* (WTO WT/DS285/R, 10 November 2004). Editors' note: In this case, Antigua and Barbuda argued that a US prohibition on the cross-border supply of gambling and betting services and certain measures restricting international money transfers and payments relating to gambling and betting services were inconsistent with US commitments under the GATS. The panel found that the US GATS Schedule did include commitments on gambling and betting services and that several federal laws (the *Wire Act*; the *Travel Act* and the *Illegal Gambling Business Act*, the latter two when read in conjunction with State laws) as well as a number of individual State laws were inconsistent with these commitments. This case had a number of important features. For the WTO,

Body substantially reversed the panel decision by upholding the US' ability to restrict Internet gambling on public morals grounds, it upheld the panel decision that the United States had included gambling in its schedule of commitments and that the ban on Internet provision constituted a ban on the remote supply of services. Some observers see potentially far-reaching systemic issues raised by this decision, with the potential to chill services negotiations.[18] Thus, some conclude that trade in services will have to continue to develop largely outside of trade rules.[19] Which in any event is happening: trade in services is finding a way as exemplified by the fact that India has become the top services out-sourcing country, not as a result of trade negotiations but as the result of market forces.[20]

This raises new questions: Will the developing countries buy into the Round? Will they provide leadership in moving the Round to a conclusion?

---

it was the first Internet-related dispute as well as the first case in which Article XIV of the GATS (general exceptions) was invoked in defence of particular measures (including, *inter alia*, for the protection of public morals and maintenance of public order—in respect of which argument the United States cited concerns about organized crime and funding of international terrorism). Further, on a more technical level, the panel decision clarified that US services commitments, absent specific alternative clarification by the United States, could be read as based on the Services Sectoral Classification List, and that the United Nations' *Central Product Classification* (CPC) breakdowns could be used to clarify the meaning of these commitments.

[18] For a commentary, see Bridges Weekly Trade News Digest, "Antigua, US Gambling Dispute: Appellate Body Issues Mixed Report", Vol. 9 No. 12 13 April 2005.

[19] Editors' note: The major alternative to multilateral liberalization is the sectoral mode, as in the telecommunications and financial services agreements. This approach may not be feasible as a general rule. Moreover, it could raise concerns among the developing countries of being a ploy to avoid Mode 4.

[20] See for example, CIOInsight, *Global Outsourcing Report*, March 2005, http://www.cioinsight.com/article2/0,1397,1776816,00.asp

The developing country perspective

Observers close to the developing countries counter that the G-20 do have real "offensive" trade interests in addition to the standard "defensive" interests and, therefore, can be "bought off" (or, perhaps better, "bought in") with the right offer. India, it was pointed out, has an offensive agenda in services, Brazil in agriculture, and China in goods.

For example, Brazil was described as having strong interests in the Round:

- There is now a big business export lobby in Brazil. Unlike the Uruguay Round, agribusiness interests are active demandeurs in Brazil's trade policy formulation.

- The WTO is the forum in which Brazil has chosen to negotiate rules; by contrast, Brazil has resisted negotiating rules in the FTAA, which resulted in the latter negotiation being pushed into an "FTAA-lite" mode.

- Brazil's offensive agenda can only be negotiated in the WTO because agriculture cannot be dealt with either in the confines of Mercosur or even the FTAA.

- Brazil and others in the G-20 are aware that they have changed the balance of power in trade negotiations and want to seize the opportunity offered by the Doha Round to exercise their new-found power.

Some thought that the negotiations would be helped significantly if the industrialized countries made a credible response to India's interest in Mode 4 (movement of services providers). However, many observers saw this as a non-starter. Mode 4 will not happen in a trade context, they argue: in the industrialized world, Ministers do not see a distinction between temporary and permanent entry; to them Mode 4 is trade-speak for immigration. That is how Mode 4 is being debated in the US Congress; the USTR has no leverage on this debate—even in the context of the US-Australia free trade agreement, temporary movement of business persons could not be included. For its part, the European Union could not extend Mode 4 to the acceding countries, and is even more worried about Turkey as a potential accession country. If Mode 4 liberalization could not

happen internally within the European Union, it was argued, it will not happen elsewhere (although it was also observed that the United Kingdom did break ranks from the rest of the European Union on Mode 4).

However, Mode 4 is important to developing countries as the importance of remittances to the balance of payments of a number of poor countries shows. Simple steps such as increased domestic regulatory transparency that might be considered as fallbacks in terms of making some progress on Mode 4 unfortunately give the real demandeurs on services such as India effectively nothing: these economies are being told, "We don't want your workers and we don't want your cross-border services".

But the Doha Round's "development" label has raised developmental expectations that have to be fulfilled somehow. A Mode 4 component may not be essential to buy-off the developing countries, it was argued; however, alternatives such as greater carve outs under the TRIPs agreement attract their own problems. Where, it was asked, is the evidence that it would be possible to expand carve-outs from TRIPs obligations for the G-90? The United States is pushing TRIPs-plus in its bilateral agreements, not TRIPS-minus. South Asia, it was noted, has signed an agreement with the United States as a precursor to an FTA but difficulties in intellectual property have slowed the process. The *quid pro quos* are hard to implement making it hard to extract market-opening concessions from the United States.

While, in terms of absolute numbers, most poor people are in the G-20, the development component of a trade deal is of particular importance to the smaller developing countries that cannot compete with G-20 big guns like China, Brazil and India. For the Doha Round to deliver tangible progress on development, it will require far more trade-related technical assistance (TRTA) than is on the table now to help G-90 Members, the countries that lack the strategic resources to develop their trade, and within that context, their broader economic development strategies. The Integrated Framework has not been impressive in this respect, it was asserted: for example, most in the G-90 still do not know how to frame their offers on services (the troubling aspect of this judgement is that, weak as the Inte-

grated Framework admittedly appears to be, it is seen by most observers as the *only* major such instrument).

Another area where development can be advanced is trade facilitation: fully one-third of the remaining gains from trade liberalization are estimated to come from trade facilitation. Asia in particular has much to gain in this area. The key to unlocking these gains is a combination of technical assistance and complementary infrastructure services liberalization. These issues are being addressed partly by unilateral initiatives and partly through regional cooperation (e.g., APEC). This element has to date been little advanced in the Doha Round negotiations but the potential is there for more—although the extent to which this contributes to a larger outcome depends importantly on the level of ambition shown in other parts of the negotiations.

A mandatory review of implementation to assure developing economies that commitments were being met would also help.

In the end-game, it was argued, the G-90 Ministers will want something concrete to buy them off. Unlike in the Uruguay Round, the hand of developing countries cannot be forced by the Single Undertaking; now that they are inside the WTO, they have a veto. Moreover, many of the smaller poor countries are better briefed and up to speed now (at least in Geneva; in capitals it is thought by some observers to be a different story). While it is not clear by how much the decision-making process has really changed, there is no assurance that the Doha Round can be concluded if there is inaction on development commitments.

Still, the greater fear is that the WTO negotiations will not deliver—which is what is providing the impetus to all the South-South trade initiatives. Thus, it was argued that the G-90 would not be a barrier to a trade agreement. At Cancún, the G-90 "walked" but that was the result of the European Union misplaying its hand by waiting too long before making concessions on the so-called "Singapore issues"[21]; in the corridors, many G-

---

[21] The Singapore issues comprised trade facilitation, transparency in government procurement, the relationship between trade and investment, and the relationship between trade and competition policy. The rubric stems from the first WTO Ministerial meeting in Singapore in 1996 at which working groups on these issues were mandated.

90 representatives were horrified at the breakdown in negotiations. The G-90 goal in the Doha Round is not to block an agreement but to shift the distributional outcomes somewhat. Their preferences are being eroded with or without their engagement in the Round; the *status quo* is therefore unsustainable.

That being said, the outcome they accept may not do much for development. The Uruguay Round was concluded at a high point of acceptance of the "Washington Consensus"; now there is reform fatigue. In the G-90, special and differential measures (S&D) is all about preserving absolute policy space and freedom from reform pressures. Indeed the July package which included S&D in various guises risks the removal of any pressures to reform.

An Alternative approach to Special and Differential Treatment.

Part of the discussion at the roundtable focussed on a proposal put forward by Andrew Charlton to move away from reciprocity as the basis for negotiations. This proposal goes into a field of other proposals for special and differential measures, such as Lamy's "Round for Free" for the developing countries, and should be considered an alternative to such other formulations.

Under the proposed modality, countries would give full market access to those that are both smaller and poorer than themselves. This would elicit a lot of south-south liberalization, especially from the middle-income countries while not exposing countries to the full force of competitive pressure from those countries that are both richer and have larger markets.

While at first blush the departure from reciprocity appeared to open insuperable problems of political economy.

- Not least, it would, in the first instance, leave the US out in the cold, as was noted immediately.
- The wide differences in competitiveness across southern countries would also be problematic: Latin America could not open up, for example, to Sri Lanka in textiles, they would face import surges.
- Such a proposal does not ensure a supply response in the poorer countries, unless combined with trade-related technical assistance.
- South-south trade is developing largely outside of the WTO. Why would they want to bring this dynamic into the WTO where it is bound?
- A distinction can be drawn between commercially meaningful liberalization and wishful thinking liberalization. The EU's "Everything But Arms" initiative falls into the wishful thinking liberalization category. Commercially meaningful liberalization must come from a microeconomic perspective, although not necessarily from simple sectoral deals. ICOA, it was suggested, is a natural experiment in commercially meaningful liberalization.

- Two key distinctions enter into the question of the feasibility of approaches to liberalization:
    - Binding versus voluntary
    - Comprehensive vs. exceptions
  It is easy to grant access except where there is tough competition. The friction points in negotiation are where the gains from trade come from. True there are safeguards; but in the current system the developing countries have natural safeguards in that bound tariffs are much higher than applied. When you get to real binding, formal safeguards are the recourse and these are much more expensive to apply and also time limited.
- And, not unimportantly, its automaticity would seem to put an important constituency for a new modality—the trade negotiators themselves—out of work.

  However, some found it refreshing to hear a new idea for approaches to trade liberalization. And, it was also noted, every visionary idea is considered impractical when first made. But, if it can be articulated in a first draft, it might make a long-term difference.

  And this proposal has some interesting features:
- We need to articulate a "fairness" principle such as this within the context of trade negotiations.
- It provides an approach to trade liberalization that naturally skews the benefits to the poor.
- Moreover, there are answers to some of the obvious objections:
    - The proposal does not exclude the possibility of the North negotiating down the MFN rates that would apply to countries that are richer and larger.
    - It could be part of a broader package to allow for cross-sector trade-offs to deal with tough issues such as agriculture and textiles and apparel.
    - Moreover, safeguards are still available to deal with import surges.
    - While it is true that there are many south-south deals are being negotiated, the costs of negotiations are very large.
    - South-south trade deals are also deals amongst non-equals. It is possible that the constraints on supply response might be weaker in the context of liberalization between developing countries than between developed and industrial countries.

Wishful thinking liberalization, it was noted, is consistent with economic thinking. For the US, it is simply a proposal to embrace free trade.

The process will matter

It was observed that the breadth of issues raised in a wide-ranging discussion on the current state of the multilateral trade negotiations can be bewildering even to some at the table; the challenges for developing countries of pulling together the various threads into a national policy in the context of a live negotiation—the domestic aggregation process—seem daunting in this light. There is accordingly some skepticism that the sort of eleventh-hour maneuvering that has typified the closing of past rounds can work in the present context.

First, there is the issue of transparency and legitimacy which in this Round has reduced the scope to strike midnight deals behind closed doors. With WTO rules reaching into internal governance issues, Ministers are now worried about internal affairs. Questions about what is happening in Geneva are being raised in the context of the domestic political process not only in the rich countries but also in the poor—in India, for example, the view is that the WTO needs to legitimize the rules negotiated in the Uruguay Round, especially in services. Simply put, trade is being politicized in the developing world—referendums are being held and there is more visibility in the press about what is being negotiated and the potential impact of particular proposals.

The G-90 was seen as more of a problem in this regard than the G-20 since within the G-90 membership is a concern about democratic legitimacy: accordingly, even if one does not believe that the WTO is about democracy in a "global governance" context, it is important to at least make a bow in that direction. For example, it was argued that Africa is not opposed to uniform global rules; the question is participation in the process by which these rules are established. Africa too has an offensive agenda in agriculture and improvements in non-agricultural market access including addressing rules of origins. But the package has to be built up carefully and the process is important.

This complicates the negotiation process. Since Seattle, there has been a shift towards a process of small group meetings (e.g., "friends of sector" meetings) with information then fan-

ning out to all Members. Members, if not direct participants, are at least represented in principle. Such a process of moving between small groups and large groups to develop and then promote proposals for ways to move the negotiations forward has been actively exploited in the agricultural negotiating group— for example, this approach was used to find a way to convert specific tariffs in agriculture to ad valorem tariffs, taking the proposal first to a mini-ministerial and then back to Geneva. But observers do not see this happening consistently in other negotiating groups. Moreover, while this approach facilitates rapid progress, it is not without its own transparency problems: mini-ministerials are, as was observed, *de facto* Green Rooms without the transparency constraints of WTO official meetings.

Second, there is the issue of the political aggregation process to support the formulation of a viable package. Members need to understand who wins and who loses from a given proposal; without that understanding, it is not possible to determine whether a given proposal is enough to buy off anyone. The African Union has a secretariat and research capacity and the ability to look at options and trade-offs and to build up positions. But the system does not work very flexibly as yet; Africa continues to face a process problem of coordinating the various regional subgroups.

Thus, there is a need for a neutral analysis of the costs and benefits to allow African WTO Members to confidently sign onto a deal. It was suggested that the OECD countries could help if they spent only five percent of their analytical resources to factor Africa into their analyses of the impacts of particular negotiating options. Recognizing that the costs are substantial, the question was raised: What is the African Development Bank doing? What about the African part of the UNCTAD staff?

Are reforms to the WTO process needed?

Squaring the need for transparency and legitimacy with the need for a manageable process raises the question: does the WTO need procedural and institutional reform?

The Sutherland Report, it was argued, would provide little help with the procedural issues, throwing in old ideas like the Consultative Group of 18 (CG18) and an annual ministerial meeting as well as reiterating what are now almost standard calls for more resources for the Secretariat and greater transparency in WTO processes.

But, as was noted, ministerials take a great deal of time and energy and can become expensive "circuses", which called this idea into question. Meanwhile, the Sutherland Report's proposal for a new consultative group is a reminder that, while the WTO still lacks a forum in which to discuss issues and still lacks sufficient analytical capacity, the proposal to use the Trade Policy Review Mechanism (TPRM) to bring out the connections was vetoed by WTO Members.

More broadly, the Report was criticized as having the hall marks of a "committee document", lacking a clear vision while rehashing well known concerns (especially the risks to the multilateral system from rampant regionalism). For the developing economies, the Report disappointed because it did not address the question, "What is the WTO all about?" Is the WTO just about trade law, or about global governance, or about development as well? Nor did the Report address the problems developing countries face in their aggregation process although it was noted that the Report's recommendation that the Director General visit the regional groups would usefully stimulate regular reports on the implications of the negotiating proposals for particular regions and countries.

But from the perspective of the way Geneva operates, the Sutherland Report recommendations *were* seen as important. Why, it was asked, is it that institutional issues only get raised towards the end of a Round? They are important and should be included at the beginning. These issues cannot be put off as suggested by some—their being on the agenda is not an excuse to spend time at the beach in Cancún or go shopping in Hong Kong! In this regard, the Report's timing was unfortunate, which contributed to its falling rather flat upon release WTO Members were, as it turned out, pre-occupied.

Towards a successful outcome at Hong Kong

To launch the Doha Round in late 2001, before completing and digesting the Uruguay Round results and China's accession to the WTO and integration into the global trading system, might well have been a great mistake. But, as one observer commented, we're stuck with it.

There are numerous sources of head winds for the Round at the present moment.

First, the multilateral trade agenda is taking place in something of a political vacuum. In particular, post-Iraq, there are questions in many observers minds about whether the United States will recommit to a multilateral agenda. And, if there is any validity to the expectation that the new US Congress will not warm up to multilateralism, what is to be done at Hong Kong?

Secondly, trade has little political traction right now in key capitals; and, to the extent it does have traction, it is often negative (e.g., resistance to Mode 4).

Thirdly, there are the various complications on the 2005 agenda, including transition in the ranks of key negotiators and the appointment of the new WTO Director General, which tend to draw energy away from the negotiations.

Most importantly, there is no clear answer to what is always the key question in a trade round: "Where are the business interests?" The political economy of trade means negotiating *through* the trade negotiators *with* the domestic interests—and the latter, except for particular narrow interest groups, are little engaged.

Do we suspend our ambitions to buy time to devise policies that work in the complex real world setting that we have? The idea of the Hong Kong WTO Ministerial locking things up is thought to be highly improbable, but a good outcome at Hong Kong does open up the road for the final run. At the same time, the expiry of US Trade Promotion Authority (in 2007) leaves little time beyond to lock things up. These considerations spark in some observers a sense of urgency: the time to involve all the Members and all the interest groups is now.

In terms of process, for Hong Kong to succeed, it was argued that it is essential that a good first approximation of the text be in hand by mid-2005; by the same token, whether or not a first approximation is in hand by that time will serve as a timely reality check for Members as to whether to start taking emergency measures. As one observer put it, officials owe it to their Ministers not to leave them with a mess to sort out. Ministers who can sit down and talk fine details with the experts are few—so it is important that the Ministerial not be burdened by too many details. The GATT Ministerial in December 1988 in Montreal was described as an abject failure for that very reason. At the same time, it is essential that Ministers participate and feel that they are participating; this can require an enormous expenditure of time and energy on the part of the Secretariat and the Chairs of the negotiating groups, but it is vital to success. At Punte del Este, it was recalled, India wanted to meet with the Chair of the Conference every evening after the conclusion of the sessions—and was accommodated. And, at the end of the Uruguay Round, Peter Sutherland supported by Friends of the Chair managed the process very skillfully. By contrast, that did not happen at Brussels and people reacted accordingly.

Given the time pressures, it was disquieting to some that there has been little progress in early 2005, although USTR Zoellick did move the markers at Davos in January, just before leaving his post, when he signaled US willingness to use different coefficients for developing vs. industrialized economies in the tariff cutting formula. The problem with losing Zoellick and EU Commissioner Lamy is that this leaves Brazil's Ambassador Celso as the only true technician among the top negotiators.

Moreover, there are all the distractions of domestic political cycles and other international agendas. For example, it was noted that the challenge to make the 2005 process work was not reduced by the delay in naming Zoellick's replacement as USTR.[22] For its part, Europe had its own internal complications: The United Kingdom, which presides over the G7-G8 in 2005 and holds the EU Presidency in the second half of the

---

[22] Robert Portman was nominated as USTR in mid-March 2005.

year, faced general elections in May which could tie up some of the energies of Britain's Peter Mandelson, who has replaced Pascal Lamy as the new EU Trade Commissioner. Then attention would switch to the internal EU process leading-up to the Gleneagles G7/8 summit. Furthermore, the Commission for Africa report to the G7 would lead to an Africa initiative at the Gleneagles summit in July. Then all Europeans would go on holiday! In September, everyone would wake up to the need to get a package together for Hong Kong—or the need to start to deflate expectations for Hong Kong.

The question to one observer thus was: do we start to talk down the level of ambition in the summer and prepare the world for the likelihood that Hong Kong will deliver a mouse? And, if so, how do we do that without undermining the WTO?

But as a tonic for pessimism, participants were reminded that there *was* a comprehensive, ambitious deal in the Uruguay Round—and no-one was more surprised than the negotiators!

Eyes Wide Shut?
Beyond Market Access in North–South Regional Trade Arrangements

Dan Ciuriak[*]

This Chapter provides a thematic summary of a seminar hosted by the Trade, Employment and Competitiveness Program Initiative of the International Development Research Centre (IDRC) in Ottawa, Canada, on 17 February 2005. The seminar addressed five themes:

- *The scope and intended benefits of North–South Regional Trade Agreements involving Asia, Africa, and Central America.*
- *Assessing the developmental benefits of regional arrangements as mechanisms for trade liberalization and regulatory convergence in non-traditional services.*
- *Do investment provisions in bilateral and regional agreements promote the development benefits of foreign direct investment (FDI)?*
- *Dealing with cross-border competition issues in the absence of multilateralism: competition provisions in international agreements.*
- *Risks and options for developing countries in negotiating new North–South agreements.*

The presentations of the seminar featured the work of members of IDRC supported trade policy research networks in developing countries, which have been investigating the challenges and opportunities posed by regional arrangements. The participants included representatives of international organizations; Canadian government ministries and agencies; IDRC staff, research partners and networks; universities; non-governmental organizations; the private sector; and experts on international trade policy and development.

A more detailed report on the seminar together with the individual presentations and papers is available on the IDRC website (see http://web.idrc.ca/en/ev-51329-201-1-DO_TOPIC.html).

[*] Deputy Chief Economist, Foreign Affairs and International Trade Canada. Comments on the text from Susan Joekes and Gerett Rusnak of the IDRC are gratefully acknowledged; responsibility for the synthesis and interpretation of seminar papers and the discussion thereof rests with the author. The views expressed here are not to be attributed to the seminar participants, to the IDRC or to Foreign Affairs and International Trade Canada.

Introduction

The rush to regionalism evidenced by the bewildering growth of regional and bilateral trade agreements (RTAs) has raised concerns that the costs and benefits for the trade system – for individual countries, and particularly for developing countries and the world's poor – have not been properly assessed. Hence the title of the seminar on which this Chapter reports: *Eyes Wide Shut?*

RTAs are generally thought to have several advantages:

- They can yield tangible benefits in the form of expanded trade more quickly than the slower and more cumbersome multilateral process.
- They can facilitate deeper and broader liberalization than might be possible in the multilateral process which now has an agenda so large that it is difficult for the membership to deal with new issues.
- They can help entrench economic reforms conducive to economic growth.
- They can give impetus to multilateral negotiations.

For developing countries that are unwilling to put their economic agendas on hold while multilateral negotiations slowly grind ahead, the regional/bilateral route is thus an attractive option. The attractiveness of the latter route for developing countries that are not yet members of the World Trade Organization (WTO) is only increased by the fact that the price of WTO accession is becoming higher: recent entrants have been required to make much greater commitments than prior applicants.

There are significant differences between different types of agreement. Southern regional arrangements, i.e. RTAs amongst developing countries alone, can generate net gains from trade, are generally believed to contribute to economic growth and can serve as platforms for, rather than impediments to, multilateral policymaking processes and outcomes. They can also facilitate harmonization of social policies and provide a forum for developing economic measures on which sovereignty can be pooled.

But they entail some risks. Most generally, RTAs compromise the underlying principle of non-discrimination at the core of the multilateral trade system. Trade rules are becoming in-

creasingly complex, with many different permutations and combinations, both within and outside the WTO. The proliferation of RTAs has created a "spaghetti bowl" of rules and commitments with well-known co-ordination and consistency problems that have frequently stalled the implementation and evolution of these arrangements in practice. Examples include the fits and starts in the history of regional arrangements in Africa, as well as in the series of agreements between the European Communities and the Caribbean, African, and Pacific countries.

RTAs spanning the North-South divide are a relatively new arrival on the scene. Economic theory suggests that they can lead to bigger economic gains than South-South arrangements because of the wider variance in the productivity and income levels between the parties. However, recent North-South agreements go far beyond providing the preferential access for exports of goods to each other's markets that generates such gains. While most now feature provisions of much wider scope than market access, there has been little systematic examination of the character or likely impact of these provisions. Thus, in their desire to gain greater market access, developing countries risk negotiating agreements without the full knowledge of what their commitments entail. Are they, to borrow film-maker Stanley Kubrick's phrase, negotiating with their "eyes wide shut"?

To appreciate the risks entailed by RTAs, developing country policymakers must gain knowledge on many issues:

- What is occurring in the trade field and in other policy domains related to globalization?
- What are the aims of the negotiating parties and the scope of the various regional agreements? Do the provisions of these agreements achieve the desired coverage and levels of commitment? What development impact do they stand to have, singly and in combination?
- Are these provisions consistent with what has been or is likely to be agreed, in the current round of WTO negotiations?

The seminar, by bringing together people from many different regions and countries, allowed a full comparative analysis of these questions, helping to build a knowledge base to help developing countries negotiate with their "eyes wide open."

North-South, South-South and Multilateral Liberalization: Perspectives from Asia, Africa and the Americas

In different parts of the developing world, regionalism is being driven by a differing mix of factors.

In East Asia, a kaleidoscope of new and sometimes overlapping preferential arrangements has emerged in recent years, including bilateral, trilateral, and plurilateral proposals that include regional as well as extra-regional partners (of which the most prominent has been the United States). This rush to regionalism can be explained in part as a response to the Asian economic and financial crisis of 1997-1998. The "contagion" that spread through the region put a new premium on regional cooperation. Even more fundamentally, given traditional economic and political rivalry between China and Japan, the tectonic shifts generated by China's economic resurgence intensified the need for a framework to minimize potential friction between these two economic powers over and above the disciplines provided for by common membership in the WTO. For East Asia, RTAs thus provide a way to advance regional economic cooperation and development objectives across a broad range of subject matters, notwithstanding the slow movement in the multilateral process, as well as a better means of achieving political objectives compared with multilateral arrangements.

South Asian countries have also been active in establishing regional and bilateral trading arrangements[1], with added impe-

---

[1] Editors' note: while the focus of the discussion on this point was South-North agreements, there have been several noteworthy developments in South Asia on the South-South front. These include the South Asia Free Trade Agreement (SAFTA), which was agreed to amongst the members of the South Asian Association for Regional Cooperation (SAARC — Bangladesh, Bhutan, India, the Maldives, Nepal, Pakistan, and Sri Lanka), and which was due to take effect January 1, 2006; the Bay of Bengal Initiative for Multi-Sectoral Technical and Economic Cooperation (linking Nepal, India, Bangladesh, Bhutan, Sri Lanka, Myanmar and Thailand); and various RTAs linking South Asian economies to China, ASEAN and other East Asian economies. An IDRC seminar in December 2005 in Hong Kong, on the margins of the WTO Ministerial Meeting, explored issues raised by these developments.

tus apparently coming from the failure of the 2003 WTO Cancun Ministerial. Typically, the key objective for South Asian countries in RTA negotiations, at least with northern partners, has been to obtain the market access they failed to obtain via WTO negotiations. However, it was argued, because these agreements often lack clear development aims, they can introduce "conditionalities" that compromise development and undermine the public interest thus reducing the benefits of preferential/duty-free market access. For example, in the case of the US–Sri Lanka negotiations, duty-free entry for Sri Lankan exports depended on Sri Lanka adopting certain policies which *inter alia* increased Sri Lanka's ready-made garment production costs and reduced affordability of drugs.[2]

A similar problem was identified with new GSP schemes being introduced by northern countries that differentiate among GSP beneficiaries. For example, the European Union's GSP-plus incentive program provides additional benefits for countries implementing certain international standards in human and labour rights, environmental protection, the fight against drugs and good governance. Many southern countries cannot meet the conditions to access these benefits, thus leading to inevitable discrimination amongst countries receiving preferential tariffs.[3]

---

[2] The measures to implement the intellectual property conditions were subsequently subject to legal challenge in Sri Lanka and were struck down by the Supreme Court; this required subsequent amendments to the implementing legislation.

[3] For example, access to the EU's GSP-plus benefits depends on ratification and implementation of some 27 conventions, 16 of which are related to human and labour rights and 11 of which deal with governance and environmental issues. While India successfully challenged an aspect of this scheme through the WTO's dispute settlement system (DS246: *EC – Conditions for the granting of tariff preferences to developing countries*), the Appellate Body opened the door to a WTO-consistent regime along these lines when it held that a GSP scheme may be considered non-discriminatory even if identical tariff treatment is not accorded to all GSP beneficiaries or if tariff preferences are addressed to a particular development, financial or trade need and are made available to all beneficiaries that share that need. See WTO News online at http://www.wto.org/english/news_e/news04_e/dsb_20apr04_e.htm, accessed December 8, 2005.

The United States has similarly tied GSP benefits to cooperation on political agendas related to drugs and terrorism. Moreover, the application of rules of origin for qualification under GSP programs can pose barriers for South Asian nations that have small domestic production bases and accordingly import production inputs; this is the case with the rules of origin applied to access the EU's GSP program. Finally, concern was expressed that RTA-stipulated obligations to adopt international standards limit the flexibilities and exceptions that WTO agreements provide to southern countries to facilitate achievement of developmental and public policy goals.

The African perspective on RTAs focussed more on economic infrastructure weaknesses that can diminish the ability of RTAs to stimulate economic growth and development. In the case of the South Africa-EU FTA, South African exports to the EU did increase, but at a disappointing annual growth rate of 4 to 5 percent, which has left the EU's share of South African exports stable at 30 to 31 percent, with little overall change in terms of either "trade deepening" or "trade widening".[4] In the case of the RTA between South Africa and the South African Development Community (SADC), categories of imports entering South Africa duty-free increased from 40 in 2000 to 70 in 2003. However, while imports from the SADC have steadily grown, growth rates in the period 1996 to 1999 were actually higher than in the period 2000 to 2003. The blame for the weak response to the FTA was laid in good measure at the door of underdeveloped transport corridors that physically limit the ability to export in great volumes as well as underdeveloped finance, telecommunications, and other services that facilitate trade relations—although problems with rules of origin and the inadequacy of domestic regulation were also mentioned.

Nowhere are differences between North–South and South–South agreements felt more keenly than in the Americas. Because of the geographical proximity of the US economy and better de-

---

[4] Editors' note: trade deepening implies exporting a narrow range of products in larger quantities while trade widening implies exporting a wider and more diverse range of products.

veloped North-South shipping infrastructure links, North-South trade dominates in terms of monetary value as well as in terms of how "seriously" it is taken, especially when trade disputes arise. There are also qualitative differences: North–South trade is based more on complementarity and comparative advantage whereas South–South trade tends to feature competition in areas of common export strength. North–South negotiations involve more issues—such as intellectual property rights, labour and environmental issues, and more broadly generate pressure to move towards a free market economy (which is often controversial, sparking anti-globalism protests); by contrast, South–South agreements mainly focus on trade.

Compared to WTO negotiations, North-South RTAs are considered less demanding in terms of negotiating resources and time; at the same time there is a greater asymmetry in bargaining power—although developing countries (even larger ones like China and India) also face asymmetric power within the WTO.

RTAs and services trade liberalization

Developing countries have been liberalizing trade in services at the unilateral, bilateral/regional (including South-South and North-South agreements) and multilateral levels. Yet progress has been slow, difficult, and usually incremental, whether in the WTO or through regional negotiations this despite the fact that economic theory suggests that services liberalization is a largely untapped source of economic growth and that there are fairly strong spillover benefits to goods-producing sectors from liberalization of producer services.

Resistance to services liberalization often is based on arguments developed to support goods protection. However, it also reflects the diversity of types of services and the complexity of the conceptual and practical issues raised in dealing with regulations and institutions touched by negotiations. In turn, this suggests the need to temper expectations of rapid progress in rule-making (i.e., in domestic regulatory reforms, emergency safeguards, and services-related subsidies) or in market-access in difficult sectors (such as air and maritime transport, audio-visual

services, education and health services, and labour mobility).[5]

Developing countries' experience with services trade liberalization leads observers to extract a number of lessons.

First, one of the more important sources of increased services trade has been unilateral liberalization, which has been achieved in good measure through privatization, demonopolization and entry by new suppliers —including foreign suppliers as foreign direct investment (FDI) policy shifted in the late 1980s from restricting inflows to actively promoting them.

Second, given the multiple tracks along which services trade is being liberalized, policy coherence is an issue. Countries need to carefully consider what commitments to make on each track, taking into account past implementation experience, both their own and that of their trading partners (including the extent to which industrialized countries have delivered on offers of technical assistance and special and differential treatment). Whereas North-South agreements, such as the Cotonou Agreement,[6] tend to demand a wider coverage of services sectors and deeper liberalization commitments than either multilateral or South-South agreements, South-South agreements tend to suffer implementation and compliance problems which diminish their impact.

Third, there is the issue of sequencing of sectors to promote development. Liberalizing producer services first improves competitiveness in other services and goods sectors, just as in

---

[5] The recent WTO dispute settlement case in respect of on-line gambling services brought by Antigua-Barbuda against the US brought out the difficulties that WTO Members can have in understanding their own commitments. In this case, the US was found to have included commitments to liberalize trade in gambling services, which according to the US was inadvertent. While the WTO decision, upon appeal, largely went in the US' favour, the fact that the US had misunderstood its own offer was telling.

[6] The Cotonou Agreement between the African, Caribbean, and Pacific (ACP) States and the EU, signed in June 2000, mandates the negotiation of economic partnership agreements (EPAs) between regional groups of ACP States on the one hand and the EU on the other. The trade components of the EPAs are to be negotiated by 2007. The EPAs are preferential trade agreements that involve progressive elimination of all trade barriers, including in services, first between countries constituting the regional EPA groups and then between these groups and the EU.

the case of tariff reductions on intermediate production inputs.

Fourth, with regard to sequencing by trading partner, full liberalization is often recommended first among developing countries and only after that with industrialized countries.

Fifth, mode 4 (movement of service providers) represents an important channel through which developing countries could gain from enhanced service market access in the industrialized world. Within developing countries, there is a strong interest in the movement of labour, which reflects the growing importance of remittances. Even liberalization to permit temporary movement could result in substantial benefits.

Sixth, supply response capacity of domestic services sectors and adjustment costs need to be taken into account when considering rapid liberalization under North-South agreements, since suppliers from industrialized countries tend to be much stronger than domestic suppliers. Even the built-in asymmetry in terms of implementing liberalization commitments under the WTO regime is not considered by some observers as adequate to address this concern.

Seventh, disciplines on subsidies are critical for developing countries, because it is difficult for them to compete with highly subsidized foreign trade partners.

Generally, the firmer the commitments on technical assistance and support for improved supply response and adjustment in the developing countries, the greater the chance that the latter will undertake the needed degree of services sector liberalization.

In terms of modalities, progress at the multilateral level in the Doha Round GATS negotiations has been minimal. However, with services aspects of RTAs only weakly constrained by multilateral rules, more ambitious, GATS-plus results can be obtained through regional/bilateral agreements.

One way is to adopt the NAFTA architecture in the RTA. Under this approach, services trade specific disciplines are established on cross-border trade in services (modes 1 and 2 of the GATS), while generic (i.e., non services-specific) disciplines are established on the movement of capital/investment (mode 3) and of service suppliers (mode 4). In contrast to the GATS' positive-list approach to specifying sectors for liberalization, the NAFTA

architecture adopts a negative-list approach, specifying the exemptions from liberalization. The latter approach has useful features: it enhances transparency by providing a complete list of non-conforming measures, both for firms and for future negotiations; and it elicits an audit of existing restrictions on international trade which is valuable from a governance perspective.

A second way is to address areas where the GATS has lagged. For example:

- government procurement of services, which has not been dealt with at all under the GATS, is an area with considerable potential for increased market access in RTAs;
- pioneering work in formulating sector-specific emergency safeguards would facilitate the liberalization of services in sectors where foreign entry can be particularly disruptive for domestic industry (e.g., the impact on retail and wholesale trade from entry of competitors like Wal-Mart); and
- the adoption of very liberal rules of origin for services and investment would prevent third party investors from bending the rules to their own ends.

A third way is to include "good governance" in RTAs:

- make commitments based on current regulatory regimes to avoid having a "wedge" between bound commitments and applied practice as is the case with tariffs on trade in goods;
- automatically reflect unilateral liberalization in bound commitments (as is done in the NAFTA);
- protect acquired rights of established service providers;
- not require commercial presence as a precondition for serving domestic markets to promote cross-border modes of services trade, including e-commerce, while engendering convergence in underlying regulatory frameworks;
- address regulatory impediments to services trade and investment (e.g., by including a "necessity" test to determine whether the restriction is unduly burdensome); and
- include transparency commitments (e.g., such as pre-notification of new regulatory measures).

Sectoral liberalization: air transport in South Africa

In the services area, sectoral liberalization tends to confront sector-specific regulatory issues and international protocols outside the scope of the multilateral trade system, as well as domestic trade-unrelated public policy objectives that can have trade-restrictive effects. Insofar as non-trade concerns tend to be shared on a regional basis, RTAs can help move the process of liberalization along. A good example of this is provided by the air transport sector in South Africa.

As general background, air transport falls within the structure of technical freedoms (including landing rights, over-flight rights, rights to pick up and drop off passengers etc.) established by the Chicago Agreement of 1944[7]. Within the framework, individual states enter into bilateral air service agreements (BASAs) that set out which commercial services may be provided within the territory of one partner by carriers from the other, such as picking up and dropping off passengers on direct flights between the two destinations. More ambitious "open skies" agreements may extend rights to drop off or pick up passengers on routes involving third countries and/or within the territory of the partner country. The air transport sector is unique unto itself because of dominant state ownership of carriers and/or a high degree of concentration; a high degree of immunity from competition; pervasive controls on entry, capacity, and pricing; and unique regimes and institutions. These measures are in place for a variety of reasons, including the fact that air transport tends to be important for regional services within a country and for national security. That being said, there has been a steady increase in liberalization within the air transport sector, with the US open-skies policy being a major driver.[8]

---

[7] *International Air Transport Agreement,* signed at Chicago, 7 December 1944.

[8] The United States is party to 60 of the 89 BASAs currently in force. In Asia, a plurilateral agreement with the United States, Chile, Singapore, Brunei, and New Zealand has replaced a set of bilateral agreements between the signa-

> Southern Africa provides a good example of the problems faced in tackling air transport liberalization. The great distances and limited transportation facilities in the region make air transport important to economic integration. The South African Development Community (SADC) constitutes a fairly significant air transport market, with Johannesburg as its hub; however, the development of the industry has been hampered by the small size of the individual national markets, high fee structures, and limited deregulation. The exit by European carriers from this regional market when the rand weakened in the early 2000s lends legitimacy to these countries' policies to ensure the participation and survival of domestically registered airlines. In this context, regionalization is thought to represent a possible way forward, although liberalization of the SADC regional market continues to encounter resistance and regionally owned and managed airlines have yet to be established.

RTAs and Investment Rules

The Uruguay Round pioneered the introduction of investment rules, in the form of the TRIMs agreement, into the multilateral framework. Since then, attempts to develop this minimalist agreement have met with little success as investment rules have been "too hot to handle". Investment was one of the Singapore Issues dropped from the Doha Development Agenda following the collapse of the WTO Ministerial meeting at Cancun. The OECD tried unsuccessfully to negotiate a "high quality" investment agreement. And the Free Trade of the Americas (FTAA) negotiations have moved into an "FTAA-lite" mode with no investment disciplines due to stiff opposition from Brazil. This reflects in part a lack of consensus not only between developing and developed but even within the developed world (e.g., US and EU positions diverge on portfolio investment, regulation in the public interest, and dispute settlement).

tories and the United States. The European Union has also been advancing the agenda with its "Horizontal Agreements" that modify existing BASAs to ensure non-discriminatory treatment of all EU members.

Progress in investment rules has largely come through bilateral investment treaties (BITs), of which there are now over 2,000 in existence. However, the emerging empirical consensus is that these investment treaties have had little effect on the global distribution of foreign direct investment (FDI).

From a development perspective, this is problematic since these agreements nonetheless do protect investment through limitations on public policy. Under BITs, for example, investors can go straight to international arbitration if they allege that the state has gone against the agreement. There has been a steady climb in the number of known investment treaty arbitrations since the mid-1990s, although the full extent of such litigation is unknown because the arbitration rules do not provide for full transparency and no notice may be given even in important cases affecting public policy. Some countries, including Canada and the US, are trying to fix this problem by requiring that all cases be made public under new agreements. However, there are hundreds of investment treaties that do not provide for transparency. Full transparency would shed light on how tribunals are interpreting agreements, which in turn would help governments understand the effect of rules which they are adopting.

One important issue raised by BITs is potential constraints on future public policy initiatives. When negotiating investment treaties, governments need to anticipate future policies and their possible relationship to investment treaty obligations. For example, in the case of state-provided health care, the expansion of publicly funded health services into new areas such as prescription drug and dental insurance could expose governments to law suits from foreign investors with a stake in those areas.

Another potential flashpoint is the issue of positive discrimination, which gives preference to certain groups. Some countries have sheltered the right to make such preferences when forming investment treaties, but most have not, leaving themselves exposed to challenges from foreign investors who allege that they face an uneven playing field. In South Africa, for example, policy tools are being used for Black economic empowerment, but some of these positive discrimination measures may conflict with investment treaties and may result in litigation.

Another issue is corporate responsibility. A Canadian NGO initiative has led to the drafting of a model international investment agreement that aims to strike a balance between investor protection and the rights and obligations of both home and host countries and also includes other elements such as corporate responsibility.[9] This sample agreement diverges quite a bit from a standard bilateral investment treaty but is generating widespread interest among developing country governments and within the private sector. One shortcoming of many BITs is that they fail to include exceptions and safeguards to shelter sensitive economic sectors like fisheries or culture, or to protect certain taxes, subsidies, or environmental regulations. As well, there tends to be little in the way of special and differential treatment written into international investment agreements. Some BITs retain more policy autonomy and flexibility than others. For example:

- some only require post-establishment national treatment whereas more advanced BITS require both pre-and post-establishment national treatment;
- some prohibit performance requirements, which are often used as a tool of industrial policy, whereas others do not mention performance requirements thus permitting them;.
- some BITs list many exceptions to national treatment and most-favoured-nation principles in economically important sectors whereas others do not.

One attempt to develop a quantitative measure of the degree of policy autonomy preserved under an international investment agreement is UNCTAD's "Flexibility for Development" index, which assigns numerical scores to agreements based on four criteria:

- the objectives of the treaty—is it pro-development?
- special and differential treatment for developing countries;

---

[9] See "Experts Call for New Rules on Business Investing Abroad," IISD Media Announcement, 27 April 2005, http://www.iisd.org/media/2005/april_27_2005.asp; see also IISD (2005), *IISD Model International Agreement on Investment for Sustainable Development Negotiators Handbook*, *http://www.iisd.org/publications/pub.aspx?id=686.*

- substantive provisions (where flexibility is valued);
- application (the force of the treaty).

With this index, statistical analytical methods can be applied to assess the development benefits of different kinds of agreements. The flexibility-for-development index should be applicable to the investment component of trade agreements.

FDI responds to a wide range of incentives, not simply investment liberalization. For example, the African Growth and Opportunity Act (AGOA), signed in 2000 and extended to 2015, has been linked to considerable foreign direct investment in a number of Southern African countries, from German auto assemblers as well in the region's apparel and textiles industry as a result of backward integration by already-present Asian apparel producers into textiles. In this case, trade liberalization in effect increased the incentives for investment. Other factors that attract FDI include the provision of public goods (e.g., raising the quality of infrastructure) and improving labour productivity (e.g., by support for training and skills development).

By the same token, greater benefits from FDI can be obtained if policies facilitate technology and skills transfer as well as the creation of linkages and networks, and if competition policy maintains space for domestic firms.

Investor protection is an important international goal, but the present system of bilateral investment treaties, it was argued, "is a stacked deck"; not enough attention has been paid to the need for these treaties to respect certain legitimate boundaries for government. It is important that governments do impact assessments to assess vulnerabilities and look at loopholes—and not wait until they are being sued.

Competition provisions in RTAs

Trade liberalization through RTAs can increase competitive pressure on domestically oriented business but it can also unleash a wave of merger and acquisition activity and an increase in corporate concentration as firms exploit new found economies of scale. At the same time, anti-competitive practices such as price-fixing schemes and abuse of dominant positions in

markets can nullify the potential gains from trade promised by RTAs.[10] Including competition provisions in RTAs is one way to improve the net benefits of these agreements. In fact, RTAs with competition provisions now number more than 100 – predominantly in North–South RTAs and mostly in the Americas and Asia.

Although there are various approaches and a wide diversity of objectives in terms of trade, competition, and development, RTA competition provisions typically create a number of obligations:

- to have a competition law and enforcement agencies;
- to have a mechanism for inter-agency notification, consultation, and co-operation; and
- to put in place supra-national mechanisms such as procedures and bodies to give effect to cooperation.

In North-South RTAs, provisions may also include undertakings regarding:

- technical assistance and transfer of best practices;
- a presumption of co-operation; and
- a presumption of enforcement priorities.

Northern and southern countries sign onto competition provisions, it was argued, for different reasons—some stated and some unstated. Northern countries tend to be interested in complementing market-opening measures and exporting their own antitrust model. Their unstated reasons may include setting precedents and fostering agency-to-agency co-operation. Southern countries seek access to technical assistance and transfer of best practices more quickly, as well as an external anchor for the entrenchment of national competition law and enforcement.

---

[10] For example, it was indicated that South-South RTAs such as COMESA, Caricom, and SADC have more than doubled trade amongst their members, while trade liberalization worldwide has increased world welfare by about $40 billion   about half of which has accrued to developing countries. But during the 1990s alone, developing countries' imports were greatly affected by price-fixing conspiracies, the welfare costs of which exceeded US$80 billion. The Central America US Free Trade Agreement (CAFTA) was cited as one example of an RTA where a lack of competition provisions may have been detrimental to its effectiveness.

Unstated motives may include fostering co-operation and raising the profile of their competition agencies and officials, and thereby raising the country's international stature. Thus, the north and south may have some overlapping motives.

Achieving desired competition results through RTAs is not easy. Competition provisions negotiated by trade policy officials may not be welcomed by the competition law officials. The value of technical assistance obtained through RTA commitments depends importantly on the quantity and type of assistance provided and the capacity to absorb. Agency-to-agency co-operation is still limited outside industrialized countries and building the trust needed for effective co-operation is a medium- to long-term process. In a similar vein, the external anchor of an international commitment is not a substitute for political commitment to competition law, including providing adequate budgetary resources and respecting the enforcement agency's independence. Compliance is important: if not all members of the agreement have a competition law in place, and those that do lack effective administrative institutions to enforce it (which it was argued was often the case), the effectiveness of the RTA will be of course weak or nil. Ultimately, the power of an RTA to deliver the expected development benefits in terms of fostering a more competitive environment depends on various other factors, which emphasizes the importance of policymakers being aware of the need for complementary measures.

There are of course viable alternatives to RTAs for accomplishing these motives, including agency-to-agency agreements and non-binding norms that may be less costly, more easily negotiated and possibly even more effective over time. At the same time, some issues, such as dealing with international cartels, are best addressed through multilateral institutions such as the WTO (although the high level of activity in introducing competition policy into RTAs has yet to spillover into the WTO which dropped competition policy from its negotiating agenda after the WTO's Cancun Ministerial). And domestic policy could contribute more to a competitive international market environment if recourse could be had in national courts for abuses corporate practices carried on abroad.

It was suggested that there is much room for additional research on the effectiveness of regional competition provisions as compared to national and multilateral solutions, the relative merits of alternative approaches and institutional arrangements, the effect of special and differential (S&D) treatment, and the experience with various issues raised by competition provisions in RTAs such as treatment of confidential information, implementation costs, and extra-territoriality of competition rulings.

Risks and options for developing countries in negotiating North–South agreements

There is a strong theoretical foundation for North–South agreements, and the potential benefits include strong complementarities between partners, technology transfer, and a sense of economic security that can lead to political security.

However, the sense of the conference was that these benefits are more modest in net terms and less easy to achieve than anticipated. When positive effects do obtain, they are often captured by vested anti-competitive interests, diminishing the development gains. There is also growing awareness of the price that has to be paid for improved market access—the need to accept rules and disciplines, the impact of which on domestic issues and development strategies might be poorly understood. Asymmetries in economic and political power between negotiating parties can skew the benefits from an RTA—an issue, it was emphasized, that is not restricted to North-South RTAs since South–South agreements are not always agreements among equals and can themselves foster new inequalities and create problems for the less developed members. Sound preparation is thus essential for a country contemplating entering into new negotiations.

Moreover, the assessment of RTAs must be made in light of the fact that agreements are increasingly overlapping as members of one regional agreement sign bilateral or regional agreements with third parties. Future negotiations conducted by a country's RTA partners can thus undermine the benefits expected from the original RTA. At the same time, smaller devel-

oping countries that are less attractive RTA partners may be further marginalized by the spread of RTAs as they are left out of the process altogether.

Thus, although trade agreements are hardly an obstacle to development goals, they cannot replace development strategies.

To some extent the sense of disillusionment with RTAs reflected a wider disillusionment about the ability of economic liberalization to trigger sustained development—indeed, in many cases, RTAs were seen as a tool to lock in other reforms. There is a growing consensus, it was suggested, that liberalization is not enough: regulation, institutions, and sound policies are also needed.

In the final analysis, the relationship between market access and development must be analyzed at the issue level—at the level of investment, competition, and services—and not just on the level of a generalised "beyond the border" agenda. For each trade issue, it was argued, it is necessary to map the provisions and exclusions that are "development-unfriendly." The assessment of risks and opportunities related to trade agreements cannot be made against an abstract idea of development. For example, while North–South agreements tend to be comprehensive in coverage and scope and address more than tariffs, they often leave out certain industries or sectors that are important to developing countries (the obvious example being agriculture). The reduction of policy autonomy from integration agreements could constrain the effective range of responses to development challenges. And environmental, labour, and other standards desired by northern countries, while potentially beneficial to developing countries, could be costly and difficult to implement while raising barriers to market access.

Conclusions

The surge in formation of RTAs has transformed the global economy. There are over 300 regional agreements notified under to the WTO (although only one to date has been pronounced as consistent with the WTO conditions for RTAs). One effect is a certain degree of fragmentation of the trading system into

trading blocs, including the European Community, NAFTA, a number of free trade agreements signed regionally with the US, and regional mechanisms in Asia and Latin America. Some observers fear that the proliferation of regional agreements is weakening the WTO because the most-favoured-nation clause is set aside in regional agreements. The question, "Are we headed in the right direction?" remains open.

WTO Dispute Settlement: Systemic Issues

Debra P. Steger[*]

On February 16, 2005, the University of Ottawa Faculty of Law convened a roundtable of leading observers of multilateral and regional dispute settlement mechanisms on systemic issues facing the World Trade Organization as it celebrates its 10th anniversary in the context of the Doha Development Agenda, the ninth multilateral round of trade negotiations but the first under WTO auspices. This note represents the Chair's thematic summary of the discussions. As these were held under Chatham House rules, no attribution is given. Responsibility for the interpretation of the discussion rests entirely with the author.

Introduction and Overview

The dispute settlement system has played a major role in the development of the WTO as an international rules-based system. Indeed, some go as far as to call the Dispute Settlement Understanding (DSU) the WTO's "crown jewel". However, while the mandated review of the DSU in 1998 and 1999 did indeed conclude that the system was working well, it nonetheless generated a large number of proposals for improvements. Dispute settlement negotiators, led by Yoichi Suzuki of Japan, had agreed to a package of reforms that would have been approved at the Seattle WTO Ministerial Meeting, had it been successful. Following the collapse of the Seattle Ministerial Con-

---

[*] Debra Steger is Professor, Faculty of Law, University of Ottawa (dsteger@uottawa.ca) and former Director of the Appellate Body Secretariat of the WTO. This paper was developed in a personal capacity; the views expressed are those of the author and not to be attributed to institutions with which she has been affiliated or to Foreign Affairs and International Trade Canada or the Government of Canada.

ference, informal consultations continued in 2000 and 2001 leading up to dispute settlement being included in the Doha Declaration which launched formal negotiations in the new Round. The negotiations on dispute settlement were to be based on the work done to date (which included issues such as the "sequencing" of Article 21.5 and Article 22 procedures, the timeframes for panel proceedings, "carousel" retaliation and third party rights) plus new issues that might be introduced by the Membership. And many new issues have been added, including various proposals to amend the panel system, to improve Appellate Body procedures, and to provide greater rights for developing countries. A still wider ranging discussion of systemic issues centred on the dispute settlement system has long been underway in the literature and on the conference circuit.

Notwithstanding the long gestation period for reform proposals, the first deadline for completion of negotiations (in 2003) passed without agreement, leaving the legacy of a failed Chairman's text[1]; negotiations have since continued, with a second deadline in May 2004 also having come and gone without agreement.

The Consultative Board to former WTO Director-General Supachai Panitchpakdi, in its recent Report on "The Future of the WTO"[2], stated that "while there are some grounds for criticism and reform of the dispute settlement system, on the whole, there exists much satisfaction with its practices and performance." However, the Consultative Board was not as positive about the "spaghetti bowl" of regional trade agreements and their relationship with the multilateral trading system. It also observed that "the WTO dispute settlement system may be at a crucial, and perhaps somewhat delicate, point in its brief history."

---

[1] TN/DS/9, 6 June 2003.

[2] Peter Sutherland et al. (2004). "The Future of the WTO: Addressing institutional challenges in the new millennium", *Report by the Consultative Board to the former Director-General Supachai Panitchpakdi*, World Trade Organization, http: www.wto.org/English thewto e/10anniv e/future wto e.pdf.

The systemic issues raised in the context of proposed reforms to the WTO dispute settlement system may be grouped into four categories:

(a) Reform of the WTO Panel System. This rubric covers a range of issues that have emerged involving the panel process, including: suggestions to establish a permanent panel body or a standing body of panel chairs to address difficulties with composing panels; the question of whether a separate Secretariat staff is needed for panels; the need for improved rules of procedure and rules of evidence; and how best to ensure independence, high quality decisions, consistency and coherence in panel reports.

(b) Approaches to Interpretation of WTO Agreements. This rubric covers a complex nexus of issues related to the question of the approach taken by the Appellate Body in interpreting WTO rules. Some commentators perceive an evolution in the Appellate Body's approach to interpretation from an essentially literalist approach in the early days, paying close attention to the legal texts with frequent explicit references to the Oxford Dictionary ("textualism"), to greater reference to public international law, relying on the Vienna Convention on the Law of Treaties as a guide to interpretation. How the Appellate Body interprets WTO rules bears importantly on various debates over its role: Does "clarifying" the rules include "filling in gaps" that negotiators left ambiguous or blank, or should it be limited to simply preserving the balance of rights and obligations as negotiated by the parties? What is the appropriate standard of review for measures imposed by, and extent of deference due to, WTO Member governments by the Appellate Body and panels? Given the very strong WTO dispute settlement system and the comparatively weak rule-making system, what is the proper balance between the dispute settlement and rule making bodies?

(c) Implementation and Compliance. This rubric covers the plethora of issues related to implementation of WTO rulings, particularly by the major powers in important cases, and the tools available to promote compliance. Retaliation

("suspension of concessions") is a blunt instrument that only the major powers can use effectively, and even then it is not always an effective and efficient mechanism. What are the alternatives, especially for smaller and developing countries? Are prospective remedies the only appropriate means? Are there alternatives in the form of damages or retrospective remedies in certain cases? What are the challenges in assessing trade damage? Given the protracted time frames over which dispute resolution unfolds, is there scope for preliminary injunctions to attenuate nullification and impairment? What can be done to shore up confidence in the system?

(d) Relationship between Regional and Multilateral Trade Agreements. In recent years, there has been a proliferation of regional trade agreements and dispute settlement mechanisms. This has raised concerns such as the possibility of conflicts of rules, "forum shopping", and double jeopardy. How serious are these risks and what are the possible solutions?

These four issue areas were explored in turn by the experts participating in the roundtable. This is an account of that discussion. In the following sections, the issues are introduced with a brief context-setting outline, followed by an annotated, thematic summary of the discussion.

The sense of the experts was that the WTO dispute resolution system has acquitted itself very well in its first decade. Some of its present challenges—including the increasing volume of cases—are for the most part the product of its successes, not its failures. Reforms, while not immediately necessary, would help to improve the system. The good news is that there does not appear to be a shortage of potentially viable proposals. An ambitious outcome to the Doha Round would likely include some procedural reforms to the dispute settlement mechanism.

Reform of the WTO Panel System and Constitution of Panels

Background and Context

Since the European Communities ("EC") first proposed the idea to establish a permanent panel body in the Doha Round dispute settlement negotiations[3], there has been considerable commentary both for and against this bold idea, and various permutations of it.

The EC originally proposed that a roster of between 15 and 24 permanent panelists be established comprised of persons with impeccable credentials and appointed on a full-time basis to ensure panelists of high quality and expertise. To avoid potential problems with conflicts of interest, the EC also proposed that these persons be required to be unaffiliated with any government, Member of the WTO, and bound by the *Rules of Conduct for the Understanding on Rules and Procedures Governing the Settlement of Disputes* ("Rules of Conduct")[4] and additional obligations like those applicable to the Appellate Body. The EC proposed an appointment process for permanent panel members similar to that used for Appellate Body Members, except that panel body members would be appointed for a single non-renewable term of six years.[5]

Because the proposal was viewed by many WTO Members as too far-reaching, the EC quickly modified its original proposal to allow the parties to a dispute to select two of the panelists from outside the panel body, within very short timeframes.[6] Thailand submitted an alternative proposal, namely the creation of a permanent panel body from which panel chairs would be selected.[7] Canada, meanwhile, recommended that a permanent panel roster be created, comprised of one person nominated

---

[3] TN/DS/W/1

[4] WT/DSB/RC/1.

[5] A proposal has also been made to amend the terms of Members of the Appellate Body to one, non-renewable term of six years.

[6] TN/DS/W/38, pp. 3-4.

[7] TN/DS/W/31, p. 2.

from each WTO Member, from which parties would be required to select panelists in individual cases.[8]

With WTO Members unable to agree on an approach, the subject of improvement of the selection of panelists did not figure in the Text prepared by the Chairman of the Special Session of the Dispute Settlement Body (the "Chairman's Text") issued in May 2003.[9]

The pros and cons of a permanent panel body

Much has been written on the pros and cons of a permanent panel body.[10] In the roundtable discussion, the advantages to

---

[8] TN/DS/W/41, pp. 3-5.

[9] TN/DS/9, 6 June 2003. In his Report to the Trade Negotiations Committee, the Chairman, Ambassador Peter Balas, noted that a number of proposals, including those relating to improved panel selection procedures, "could not be included in the Chair's proposal in the absence of a sufficiently high level of support…"

[10] In support of its proposal, the EC cited the growing number, complexity and duration of WTO cases, compared with the previous GATT system and the increasing difficulty to select qualified panelists which has led to significant delays in the panel process. The EC also suggested that a permanent panel body would: be better able to handle new procedural developments, such as preliminary rulings and special rules governing the treatment of business confidential information; result in faster proceedings and produce better and more consistent rulings, with the result that fewer panel reports would be reversed by the Appellate Body; increase the legitimacy and credibility of the panel system in public perception; and increase the involvement of developing countries in the panel selection process and would allow more of their nationals to be appointed to panels. TN/DS/W/1, at pp. 2-4. For a discussion of the permanent panel body concept, see William J. Davey, "The Case for a Permanent Panel Body", in *Part IV: Grouping on WTO DSU Reform: Proposals Relating to First-Level Permanent Panels*, 6:1 *Journal of International Economic Law* 175-186 (March 2003). Professor Davey sees the establishment of a permanent panel body as an "inevitable and natural outgrowth of the Appellate Body" (at p. 185). For an earlier article, see William J. Davey, "A Permanent Panel Body for WTO Dispute Settlement: Desirable or Practical", in Daniel L.M. Kennedy and James D. Southwick (eds.), *The Political Economy of International Trade Law: Essays in Honor of Professor Robert E. Hudec* (Cambridge: Cambridge University Press 2002) 496-527. See also: Debra P. Steger, "WTO Dispute Settlement: The

moving to a permanent panel body (PPB) were summarized as follows:

(1) There would be significant time savings in organizing panels. At present, settling on the composition of a panel can take two months, while arranging the meeting schedule can take an additional one or two months given inherent difficulties in coordinating divergent schedules of a group of *ad hoc* panelists who are frequently geographically dispersed and busy with other work.

(2) More experienced panelists are required to cope with the increasingly heavy and complex workload under the WTO's Dispute Settlement Understanding compared to the GATT era. This reflects, *inter alia*: the introduction of whole new areas of WTO substantive legal competence, such as trade in services and intellectual property; the increase in complexity of rules and complications raised by reliance on confidential business information in areas such as anti-dumping, safeguards and subsidies; the complexity of factual issues raised in cases dealing with sanitary and phytosanitary (SPS) measures, technical barriers to trade, and other health and environmental issues; and the general increase in complexity of claims with the progressive legalization of the process. In this context, the current uneven level of experience of panelists (many of whom serve on only one or two panels) translates into an uneven quality in panel reports, which in turn increases the frequency of recourse to, and subsequent reversals by, the Appellate Body.

(3) A permanent panel body, through more frequent interaction and greater legal expertise of its members would be better placed to standardize panel working procedures and develop rules of evidence. At present, procedural rulings are developed as *ad hoc* decisions with Secretariat guidance. Moreover, a permanent panel body would be better able to respond to increasing demands for procedural innovation in areas such as preliminary rulings on difficult issues and

Next Step", in Debra P. Steger, *Peace Through Trade: Building the WTO*. (Cameron May, 2004): 327-346.

provisional relief; and would be better able to efficiently deal with post panel proceedings (remands, Article 21.5 compliance proceedings, Article 22.6 arbitrations, etc.).

(4) A permanent panel body would increase the legitimacy of panel decisions by eliminating or at least reducing the appearance of conflict of interest raised by the practice of staffing panels with government officials many of whom are delegates of missions in Geneva.

(5) A permanent panel body would likely result in a better distribution of the work on panels among the WTO membership than has proven possible under the present system, including ensuring more frequent participation by individuals from developing countries and also greater balance in representation of the largest Members.

(6) A permanent panel body would likely result in a reduction of the influence of the WTO Secretariat. While the support of panels by the Secretariat has improved the quality of panel reports, it has also led to the perception that the Secretariat has too much influence on decisions made at the panel level. Since a permanent panel body would ultimately have its own supporting secretariat, the main difference would be in the separation of the function of advising panels from the functions of advising the WTO Director-General and Members, a separation that would eliminate any suspicion that Secretariat staff were trying to implement their own views on interpretation of the agreements through their work with panels.

However, several potential risks in, and objections to, moving to a permanent panel body were also identified.

First, there is issue of political participation in the panel selection process.[11] Governments, it was observed, jealously guard

---

[11] See, however, the contribution by Chile and the United States to the DSU negotiations in which they proposed a number of options for "improving flexibility and Member control in WTO dispute settlement". TN DS W 28. Andrew Shoyer shares these views. He argues that the current *ad hoc* selection system "serves to address one of the key objectives of a party to a WTO panel proceeding - to compose a panel that appears likely to rule in its favor - better than would a permanent panel body." Andrew

their prerogatives. Moreover, "politicization" of the selection process by the WTO is perhaps the biggest worry within the WTO Membership ("The WTO can't be trusted to appoint responsibly"). Some argue, however, that this concern is made largely moot by the fact that, since the Members cannot usually agree on a slate of panelists for particular cases, the Director-General ends up appointing most panels. In effect, party control no longer exists in most cases anyway because of the problems of composing panels.[12]

A related problem with a Panel Body, it was pointed out, is its acceptability more broadly, not just to governments but also to civil society. The WTO is seen as part of a system of global governance and decisions by the select group of panelists could have potentially far reaching consequences.

Second, it would not be easy to achieve geographic balance, including developing country representation, in a small group, not least because lawyers from the United States and the EC would be well represented. By the same token, since these two Members are frequent litigants, it would seem likely that the present rule excluding nationals from the parties to the dispute from sitting on a panel would have to be eliminated.

Third, there would be a loss of case- and agreement-specific expertise that resides in the larger pool of panelists; there is some reality to this risk since there is a question as to how quickly the Permanent Panel Body would gain experience.[13]

Shoyer, "Panel Selection in WTO Dispute Settlement Proceedings" in *Part IV: Grouping on WTO DSU Reform: Proposals Relating to First-Level Permanent Panels*, 6:1 J Int'l Econ L 203-209 (March 2003); at p. 208.

[12] As William Davey has pointed out, "The selection process has become more difficult over time — whether because of an imbalance between the demand for and supply of suitable panelists (complicated by the increased time commitment now expected of panelists compared to the past) or because of increased party selectivity as the stakes of dispute settlement have become more significant." William J. Davey, " The Case for a WTO Permanent Panel Body", *op. cit.*

[13] For example, Andrew Shoyer contends that the current system can facilitate panel composition of "a higher quality" than a permanent panel body because it draws heavily from active trade policy practitioners who are

A related question is: what would be meant by qualifying "experience"? Only GATT/WTO experience? Or should broader experience of a relevant sort also serve to qualify individuals for the permanent panel body—for example, prior participation on other international tribunals and domestic courts?

Fourth, there is the concern of potentially excessive costs in maintaining the Permanent Panel Body. The main cost would be retainers to ensure that panelists put the necessary priority on panel business (it is disruptive to panels when individual panelists fix dates to return home to attend to their other business, even if the panel business has not been wrapped up), and indeed that they allocate the necessary time (otherwise, the panels become dependent on the WTO Secretariat). However, based on a notional scoping out of the likely costs, it was suggested that this does not appear to be a serious argument.

Other proposals for improving procedures for panel composition were also debated.

One option, along the lines suggested by Canada, would be to establish a proper panel roster. The current indicative list is not generally used. The question was raised: could current practices and attitudes toward qualifications for panelists be changed?

Another option would be a version of Thailand's proposal to establish a standing body from which panel chairs would be selected.[14] This approach would provide some of the advantages

---

employees of governments in capitals and in Geneva. echoes the views of some WTO Members concerned about losing "Member control" over the system. Andrew Shoyer, "Panel Selection in WTO Dispute Settlement Proceedings" *op. cit.*

[14] A similar proposal has also been made by Professor Thomas Cottier. Concerned that a Permanent Panel Body would upset the balance of power between the dispute settlement system and the political bodies of the WTO and disrupt the traditions of GATT/WTO dispute settlement, he supports a more evolutionary approach. This would involve the establishment of a College of Permanent Panelists composed of a defined number of alternating panelists and a small group of full-time or half-time chairs for panels who would reside in Geneva. The College would be supported by an independent legal and clerical staff like the staff of the Appellate Body. Thomas Cottier, "The Permanent Panel Body: A Bridge Too Far?" in *Part IV: Grouping on*

of a Permanent Panel Body, for example in terms of improved expertise and procedural standardization. However, it would not be as effective as the Permanent Panel Body system in terms of time saving and flexibility advantages. Moreover, it was noted, frictions would develop with a Panel Chair Body due to the inherent inequality that would obtain with the Chairs being much closer to the Secretariat than the other panel members.

A third option would be along the lines of the EC's modified proposal to establish a Permanent Panel Body but to include flexibility to appoint, on a case-by-case basis, one or two outside panelists. This approach would arguably retain most of the efficiencies of the Permanent Panel Body (depending on the frequency of recourse to outside panelists), while also allowing for specific expertise to be drawn on for particular cases and providing a partial safeguard against "bad" appointments to the Permanent Panel Body. A variant of this approach would be to provide for at least partial rotation in order to provide possibility for individuals from all WTO Members to participate.

While the Permanent Panel Body concept was seen as generally a good idea, and some of the risks and concerns were seen as less real than perceived, the sense of the roundtable was that a Permanent Panel Body was unlikely to be established any time soon. WTO Members simply are not ready to take that big a step.

This resistance is not necessarily a mystery, as one observer noted. Nonetheless, it is a paradox that WTO Members accept rulings from a professional Appellate Body, but resist having a permanent tribunal at the panel stage. The background suggests that Members, overwhelmed by the prospective demands of implementing the Uruguay Round agenda, sought to strengthen the Dispute Settlement Understanding as a way to safeguard their interests. The Appellate Body was accepted on this basis. Not much activity was expected; Members were therefore surprised by the level of activity and what some perceive as the "activism" of the Appellate Body.

Gradualism rules in the WTO: GATT/WTO reforms have historically been slow, usually preceded by provisional rules, developed from practice with cases. Accordingly, a cautious approach is more likely to be taken to further reform the system.

Provision of Secretariat Advice to Panels

The WTO dispute settlement system is not really a panel system, it was noted, but a panel and a secretariat system. Various problems have been identified in the current structure under which the WTO Secretariat provides advice to panels.

The Secretariat is organized to support the WTO both in negotiations and dispute settlement. As a result, there can be conflicting demands, which requires a large number of players. On the one hand, this hinders standardization of advice provided to panels; on the other hand, it creates a perception in the Membership of excessive power being wielded by the Secretariat.

One option to address the problems of standardization and timeliness of support would be to create a separate staff for dispute settlement panels, along the lines of the independent Secretariat of the Appellate Body. This would be inevitable with the Permanent Panel Body (either provided for at the time of its creation or soon thereafter upon the demand of its members). The question arises whether this approach would be feasible with *ad hoc* panelists?

One option to address concerns within the Membership about possible overweening influence by the Secretariat would be to require it to circulate to the parties its advice to panels. However, this approach might simply generate additional procedural complexity without effectively sidelining the concerns (since, for example, any such rule could be easily circumvented by oral advice). Also, a requirement that the Secretariat make its advice to panels available to the parties may hinder the close, confidential working relationship between panelists and Secretariat staff, and make panel deliberations more difficult.

Rules of Procedure and Evidence

There is room for improvement with respect to rules of procedure and evidence.

While the procedural rules under which panels operate have standardized informally, variations are inevitable with *ad hoc* panelists. This has resulted in procedural rules remaining a recurring issue within the WTO.

Similarly, with respect to evidentiary rules, while most cases do not raise major evidentiary issues, there is limited experience in the ad hoc panel system for dealing with such issues effectively and appropriately when they do arise. One particular problem that needs to be addressed is how to deal with business confidential information.

Many of the perceived problems with panels in these regards could be addressed by professional training for potential panelists. Such training, which is routinely provided to judges and members of administrative tribunals and addresses issues such as rules of evidence, procedure, role of experts etc., could be led by experienced judges and jurists. Such training would also provide an opportunity for panelists to get together and "bond" which would improve their ability to work together.

And, as one observer noted, some of the discontent with the quality of decisions might usefully be aimed at the quality of the advocacy: if the quality of the advocacy is indeed inadequate, panels are provided poor material to work with in fashioning their decisions.

Is reform possible?

While not all are convinced that the system is broken and needs fixing, the Membership and the trade policy community at large see a potential benefit from some improvements to the panel process.

The key problem may lie in the fact that panel reform is not easily separated from the issues of reform of the political decision making bodies of the WTO and redefining the role of the Secretariat. In other words, panel reform inevitably will raise

fundamental issues related to the separation of powers—and ultimately the balance of powers—between the so-called "legislative" and "judicial" arms of the WTO in the emerging WTO constitutional system.

Approaches to Interpretation

The interpretative choices made by the Appellate Body are of broad systemic significance. A particularly important policy decision of the Appellate Body has been its practice, followed in its earliest decisions, of "textualism"—discerning the "ordinary meaning" of the words of the treaty in light of their object, purpose and context—to inform its interpretation of the WTO texts rather than the negotiating history.

This policy decision has a firm foundation in the customary rules of treaty interpretation, as reflected in Vienna Convention Articles 31 and 32. These Articles allow resort to sources such as preparatory work and the negotiating history only under circumscribed conditions:

(a) where the meaning of a treaty provision is ambiguous when interpreted in light of its object, purpose and context, as well as with the aid of other, relevant international legal rules and instruments;

(b) to confirm an interpretation pursuant to the rules in Article 32 of the Vienna Convention; or

(c) to avoid an otherwise absurd reading of a treaty provision.

Fair enough. But, as some observers noted, ordinary meaning is not decisive, since the context in which words are used matters. Context is particularly important, it was argued, with respect to the WTO Agreement because of "constructive ambiguity" in the framing of the legal texts. What can be made by the interpreting body of "ordinary meaning" without the contextual information provided by the preparatory work and debate which lays bare the conflicting interpretations papered over by purposeful ambiguity?

By the same token, others point out, the negotiators' actual intent is reflected only in the final text of the Agreement; Members have declined to take up the opportunity afforded by the

WTO Agreement to put forward clarifying statements on issues they were unable to settle in negotiations, and there is little help from other quarters. The text accordingly must be the primary focus for the interpreting body.

This leads to the conundrum of whether the Appellate Body should give a decision on an issue on which the Members could not agree in negotiations. Some think not, seeing this as "judicial activism".

Others point out that the Appellate Body *must* rule on whatever is thrown at it. The WTO Dispute Settlement Understanding provides for compulsory jurisdiction in dispute settlement cases, and explicitly requires the Appellate Body to address the legal claims presented to it. The fact that the Appellate Body must adjudicate saves it from speculation about its motives for taking on a case or not. The parties might consider a decision not to adjudicate to be politically motivated which would undermine the system.

To some, there is less here than meets the eye; the Vienna Convention, it was noted, combines in one clause the various interpretative approaches of the ordinary meaning of the words, the context, and the object and purpose of the treaty. The relevant test by this view is legitimacy which depends on the quality of the decisions and their broad acceptability—which to some practitioners depends in turn at least in part on the consistency and predictability of the Appellate Body rulings.

In this regard, there is a big difference between interpreting statutes and interpreting treaties because in the latter case the interpreting body is dealing with sovereign States. In domestic courts, judges are skeptical about assertions and want proof. By contrast, panelists face States and must defer to their assertions of fact on evidentiary issues. There is accordingly less of the careful sifting and weighing of evidence in WTO litigation than in domestic law. And much depends on how the parties plead the facts.

What then explains the choice of the Appellate Body to take on political issues sometimes, but not other times?

It has been said in this connection that "the fate of the Appellate Body depends on the personalities". This, it was argued,

is partly true: early interpretations such as *US - Shrimp* reflected the broad backgrounds of the originals Members of the Appellate Body, their legal training and experience, and their political sensibilities as they examined the treaty text and wove their way around the landmines. Similarly, in the *India-EC GSP* challenge, which was based on the "non-discrimination" requirement in the GATT GSP provisions, had the Appellate Body upheld the panel decision, it would have impacted the GSP system. And it chose not to.

But this view of the Appellate Body is also partly incorrect: it was argued that, on the record, the Appellate Body sorts out the key "A" cases, for which it mobilizes the heavy duty intellectual power, from the secondary "B" cases, in which it applies a more simplistic, textual approach.

Accordingly, one has to look at the Appellate Body's reasoning in terms of what Dworkin called "fit". Interpretation may begin with the text itself, but nevertheless requires the interpreter to introduce a significant amount of outside material such as the interpreter's intuitions about the treaty's provisions and the meanings of its words in light of the world which generates them. For example, in addressing the "like products" issue in *EC - Asbestos*, the Appellate Body imagined the way in which consumers might distinguish between products in light of their health effects, assuming a world where liability rules would internalize negative health externalities of products.

While the Appellate Body Membership has changed, it has tended to follow the interpretative pattern of early Appellate Body decisions, a pattern which, it was argued, reflects sound and perspicacious intuitions about the challenges posed by the cases in question.

Implementation and Compliance

The range of issues that surround implementation of WTO rulings is long.

Non-compliance was a problem in the GATT system which preceded the WTO. As a result, there were persistent attempts to strengthen multilateral dispute settlement, which include the

1979 amendments in the context of the Tokyo Round, the 1989 *Dispute Settlement Procedures Improvements* which closed the loophole that a Member could block the establishment of a panel, and the WTO *Understanding on Rules and Procedures Governing the Settlement of Disputes* which made adoption of panel and Appellate Body reports and authorization of retaliation more or less automatic.

Overall, the WTO dispute settlement system is working well. Unilateralism, which was a significant problem before the conclusion of the Uruguay Round, has been kept at bay, and Members have made more use of the dispute settlement system than ever before. Since 1995, there have been well over 340 disputes brought to the WTO, leading to around 90 panel reports and 65 Appellate Body reports. Approximately one-quarter of all cases settle in the diplomatic phase, without recourse to a panel or arbitration body. While the United States and the European Communities remain the major users of the dispute settlement system, developing countries (especially the larger and more powerful among them, such as Brazil, India, China and Mexico) have become more active players in the system both as complainants as well as defendants. Many less developed countries, however, have never participated as a party or a third party in a case. While the record on compliance is good—80 percent of the cases are implemented within a reasonable period of time—there are, however, problems in a handful of major, politically-sensitive cases that test the limits of the system.

Discussions on reform of the WTO dispute settlement system have been taking place, off and on, since the 1998-99 review process. Since the beginning of the Doha Round, there has been a myriad of proposals by Members, ranging from major reforms (such as the EC proposal to establish a Permanent Panel Body) to smaller procedural reforms. The Chairman of the Dispute Settlement Body issued a "Chairman's Text" in 2003 that set out proposals that he considered had some measure of support from Members. Work on this subject has since continued in informal groups in Geneva. While the dispute settlement negotiations were not originally linked to the Doha Round's single

undertaking agenda on the assumption that there could be an early conclusion on these systemic issues, it now appears that, if there is to be an outcome on dispute settlement, it will likely coincide with the conclusion of the Round.

Are reforms to the dispute settlement system necessary? The answer to this question is: "no". The system can be, and has been, improved through experience and practice in cases. Examples include: sequencing between resort to compliance panels under Article 21.5 and requests for authorization to suspend concessions under Article 22; and the more controversial experience with treatment of *amicus curiae* briefs. While most WTO Members remain steadfastly opposed to the idea of opening up panel meetings to observation by the public, in the current *EC – Hormones* case, the parties (the United States, the European Communities and Canada) agreed to open up the panel meetings.

It is in the area of compliance/implementation that most of the Members' attention should be focused. Suspension of concessions or retaliation is a blunt instrument that realistically only the two most powerful Members, the United States and the European Communities, can use effectively; most Members recognize that other means need to be found to encourage Members to implement when encouragement is needed. Also, there are certain types of cases—for instance, antidumping, countervail, safeguards and government procurement—in which, given the timeframes for WTO dispute settlement, there is often no effective remedy at the end of a dispute because, under the WTO, there is no provision for retrospective remedies or preliminary injunctions to prevent injury from occurring while the case is being heard. Several interesting proposals have been made by WTO Members on compliance/implementation, including the idea of providing for monetary compensation for damages. These proposals are worthy of serious consideration in order to make the WTO dispute settlement system more effective and predictable.

A proposal that seems to be developing some resonance with Members is the idea to provide the Appellate Body with the power to remand an issue involving findings of fact back to

the panel. This would be helpful in cases in which the Appellate Body has reversed the panel on a legal point, but finds it cannot complete the analysis. In most cases, such rulings are not fundamental to resolving the dispute between the parties, but there may be some cases in the future which are left unresolved if the Appellate Body is not given either the power to remand a case to a panel or to examine questions of fact upon appeal.

Observers agree that professionalization of the panel level through the creation of a permanent panel body or roster is the logical next step in the evolution of WTO dispute settlement. However, there is not much support for this idea among WTO Members, and it is not likely to be a major issue in the final stages of the Doha Round negotiations.

Technical assistance to developing countries involved in dispute settlement activities continues to be a major concern for most Members. The Advisory Centre on WTO Law is a significant addition to the capacity of developing countries engaged in dispute settlement, but clearly, more needs to be done. Some innovative proposals have been made by developing countries, including the least-developed among them, in this Round; these proposals deserve careful consideration. The idea of awarding costs to successful developing country complainants is a particularly interesting proposal.

Retaliation and its alternatives

Under the WTO *Dispute Settlement Understanding*, enforcement falls to the complaining party or parties, even when the case involves systemic issues. There is no provision for collective action; this framework thus takes no account of the capacity of the complainant(s) to retaliate effectively.

This places smaller economies in a more difficult position than larger economies. Small economies have little leverage and, being dependent on imports as well as exports, suffer the self-inflicted pain of retaliation more deeply than larger, diversified economies with greater capacity for import substitution. But even the major powers that can leverage compliance through their capacity to retaliate have proved to be reluctant to

actually exercise this right; as a matter of fact, retaliation has only rarely been used.

Further, the present framework for retaliation only deals with prospective remedies and does not address reparation of trade damages, which in the view of some should be the real issue.[15] (As one observer noted, developing countries were shocked to find that the end result of litigation in the WTO was not cash!)

And, retaliation is a blunt instrument. Retaliatory trade measures hurt consumers or industries in the importing country that were not involved in the dispute and may target private producers in the exporting country that were innocent bystanders to the dispute.

This leads to a series of difficult questions:

- Are prospective remedies adequate and appropriate for all of the WTO agreements?
- Are retrospective remedies and calculation of *actual* trade damages possible alternatives?
- Should remedies provide reparation as well as an inducement to comply?
- Is injunctive or interim relief a possible remedy?
- Should different remedies be designed to fit the needs of specific agreements?

Another alternative is Robert Lawrence's concept of Contingent Liberalization Commitments (CLCs). Under the CLC scheme, WTO Members would stipulate a pre-authorized compensation mechanism during negotiations. A country that has suffered nullification or impairment in one area could choose from the list of CLCs on standing offer from the infringing country to obtain an appropriate degree of offsetting benefits. This approach maintains reciprocity while providing incentives for compliance, but does so in a liberalizing, not trade-restricting and self-damaging manner. Rather than a "shot in the foot", CLCs would be a "shot in the arm".

---

[15] See for example, Marco Bronckers and Naboth van den Broek, "Financial Compensation in the WTO: Improving the Remedies of WTO Dispute Settlement", *Journal of International Economic Law* 8(1), 101–126

However, Lawrence's proposal is still open to the criticism that the firms and persons who are actually hurt by the infringements are not the ones who would necessarily receive the offsetting benefits in either protection through retaliatory tariffs or additional market access through CLCs.

The advantage of compensation is that it is financed by the entire tax base of a Member; this addresses, to some extent, the fairness issue in terms of spreading the pain within the country maintaining the non-complying measure; it leaves open the question of fairness in the country receiving the compensation.

The economics of retaliation

There are also the significant difficulties of assessing damage from non-conforming measures. These go well beyond the question of "Where do the numbers come from?"—although that question itself elicits troubling answers (one sample answer: "They come from out of a hat!"). While there is a need for stronger economic analytical capacity in the WTO Secretariat, the more deep-seated problem is that the results of economic modeling are considered by some to be so speculative that they have little credibility.

There is a further question of the effectiveness of the remedy. Economists are questioning whether the DSU is as strong as the legal community makes it out to be. The way markets work undermines the capacity of retaliation to inflict pain. In commodities or commoditized manufactures traded on wholesale markets, a change in tariffs on shipments going to one country is easily circumvented. Also individual companies will hop tariff walls with investment—e.g., by buying out the company protected by the tariff.

These considerations suggest that there is a need for more cross-pollination of ideas between lawyers and economists, not only in the administering the dispute settlement cases but also in helping to shape the reforms to the DSU itself.

But there legal problems as well. For one thing, there is a question as to whether the presumption of nullification or impairment can ever be rebutted. DSU Article 3.8 provides that, in cases where a panel or the Appellate Body has found an infringement of the obligations under a covered agreement, the government measure is presumed *prima facie* to constitute a case of nullification or impairment. This means that there is normally a presumption that a breach of the rules has an adverse impact on other Members party to that covered agreement; in such cases, it is up to the Member against whom the complaint has been brought to rebut the presumption. Nevertheless, WTO panels and the Appellate Body have generally followed the 1988 *US - Superfund* GATT panel ruling, which suggested that this presumption is irrefutable.

If so, has this presumption precluded a fair assessment of trade damage by arbitrators? In an arbitration under DSU Article 22.6 to determine the level of nullification or impairment in *US-1916 Act*, the arbitrators concluded: "the original panel in the dispute found, and the Appellate Body confirmed, that 'the 1916 Act nullifies and impairs benefits accruing to the European Communities.' Therefore, while the level of nullification or impairment has not been specified in quantitative terms in the EC request under Article 22.2, it clearly is not, and cannot be, 'zero'." (para. 5.48)

When and how should the "level of nullification or impairment" be assessed? Was it wise to postpone the assessment of "adverse impact" until the point at which the complainant requests authorization to suspend concessions? If a panel were to assess trade damage when it determines liability, would that provide a more powerful inducement for the defending party to comply? Under the current system, where remedies are prospective, the assessment of nullification or impairment does not cover damage incurred prior to the end of the "reasonable period of time" given to the defending party to comply. Should it cover such actual damages; for example, by starting the liability clock ticking at the date of filing the case? Or at the date of

adoption of the panel and/or Appellate Body report by the Dispute Settlement Body?

Should the Appellate Body play a role in the arbitrations on the calculation of damages? There have been seven disputes in which the level of nullification or impairment was determined by the panel in an arbitration under DSU Article 22.6 (*Bananas, Hormones, Brazil Aircraft, Canada Aircraft, FSC, 1916 Act, Byrd Amendment*). Should these arbitral awards have been subject to review by the Appellate Body?

Relationship between Regional and WTO Systems

Regional trade agreements (RTAs) are proliferating, generating a complex web of multilateral, plurilateral and bilateral treaty obligations. There is a wide range of opinion on this trend, from "what's the fuss" to "this is damaging the multilateral system". The sanguine view appears largely based on the provisional consensus that RTAs on balance have been trade liberalizing, creating more trade than they have diverted, and have usefully pioneered ways to advance liberalization in new areas.

However, the proliferation of RTAs has also raised concern over the erosion of the two pillars of the WTO: the non-discrimination principles of most favoured nation (MFN) and national treatment.[16] Less than one-third of merchandise trade is now carried on under the MFN clause.

Moreover, the proliferation of RTAs and the increasing fragmentation of the global trading system seem impossible to stop or even to slow down. To be sure, the consequences of fragmentation are not too serious yet; moreover, RTAs come in batches, with considerable similarity of provisions within each batch. Nonetheless, there is concern that the more ambitious

---

[16] See, for example, Chapter 2, "The Erosion of non-discrimination" in Peter Sutherland et al. (2004). "The Future of the WTO: Addressing institutional challenges in the new millennium", *Report by the Consultative Board to the former Director-General Supachai Panitchpakdi*, World Trade Organization, http: www.wto.org/English/thewto e 10anniv e future wto e.pdf

second generation agreements now being negotiated will lead to greater problems in the coming decades.

Thirdly, there is concern over the unsettled legal status of RTAs. GATT Article XXIV and GATS Article V expressly allow RTAs, but only if they meet certain legal requirements. However, consensus on whether an RTA conforms to GATT Article XXIV appears only to have been achieved in one case in GATT/WTO history.[17]

There is great potential for conflicting obligations with the complex web of RTAs; how they relate to the WTO Agreement has become a major cause for concern. This is complicated by the fact that RTAs have their own dispute resolution mechanisms and disputes can be brought under both RTA and WTO mechanisms.

One much discussed issue is "forum shopping". This is a real issue: in one anti-dumping action, Brazil took Argentina to the WTO for one part of the dispute and to a tribunal under MERCOSUR for the other part. That being said, the WTO is often the forum of choice—there are indications that parties go to the WTO because the regional mechanisms are not as effective or binding.

A second, rather more subtle issue is that, while the RTA mechanisms acknowledge the WTO DSU and its provisions (e.g., WTO decisions/interpretations have been incorporated in regional agreements such as MERCOSUR, the WTO Agreement, with some exceptions, does not explicitly recognize the RTA provisions.

- For example, the dispute in *Canada - Periodicals* was difficult for the United States to take to NAFTA, because of the cultural exception in that agreement, so the United States went to the WTO where no such provisions exist. Should Canada have been able to plead the NAFTA provisions in

---

[17] The Czech-Slovak customs union was found to be consistent with WTO rules. See Jo-Ann Crawford and Sam Laird, "Regional Trade Agreements and the WTO", Research Paper No. 00/3, Centre for Research in Economic Development and International Trade, University of Nottingham, http://www.tradeobservatory.org/library.cfm?refID=25329.

the WTO? In the *Canada - Autos* case, Canada pleaded the Canada-US Autopact provisions, which were incorporated into the NAFTA, in its defence, but the panel did not apply them in deciding that dispute.

▪ WTO law is important for investor-state disputes, but there is no movement of ideas in the opposite direction.

▪ Meanwhile, the European Court of Justice, which is part of a regional agreement, refuses to apply WTO law.

Some observers suggest that perhaps there ought to be a dialogue rather than the present one-way interaction between the different jurisdictions.

The proliferation of bilateral investment treaties (BITs) is also an important development with potential systemic consequences.

Concluding thoughts: the glass is more than half full...

While reform of the WTO dispute settlement system is not on the front burner as critical to the success of the Doha Round, there will be an opportunity in the final stages of the negotiations to make some improvements. The scope for reform is likely to be limited, however, given the major changes made during the Uruguay Round. Also, the dispute settlement system is capable of incremental reform through practice and experience in individual cases. If negotiators were to focus on one area for reform of the dispute settlement system, they should seriously consider improving the rules and procedures relating to compliance and implementation of DSB rulings and recommendations. Several useful suggestions have been made by Members in this area that are worthy of serious consideration to make the WTO dispute settlement system even more effective and predictable.

The dispute settlement system of the WTO will continue to consider disputes and clarify the rules, whether or not the Doha Round is successful. The Doha Round provides a unique opportunity for Members to consider making improvements to the dispute settlement system. The functioning of the dispute settlement system has also drawn attention to the need for negotia-

tors to be clear in expressing their intentions when drafting new agreements or modifications to the existing agreements.

The proposals in the Sutherland Report on the functioning of the WTO system also deserve serious consideration by Members to alleviate concerns about the legitimacy of the WTO and to make it a more accountable and more transparent international organization.

What the Appellate Body Wants

Ambassador Julio A. Lacarte[*]

It is an impossible presumption to suppose what the Appellate Body might want—especially from one who is no longer a Member of it. Yet the question has been put to me, and it deserves a fair answer in the interests of sharing with trade practitioners any insights from my own past experience serving as the first Chair of the Appellate Body and, hopefully, in the interests of facilitating the work of my successors in the future. So without further apology, what might the Appellate Body want to see in terms of developments in the multilateral trading system?

A Successful and Ambitious Outcome in the Doha Round

First and foremost, it is in the interests of the Appellate Body and of the international trading system generally that there be a successful outcome to the Doha Round of multilateral trade negotiations.

The World Trade Organization (WTO) will face a possible crisis, and its dispute settlement mechanism will be put under tremendous strain, if the Doha Round fails. With the expiry of the so-called "peace clause" in the WTO Agreement on Agricul-

[*]Ambassador Julio A. Lacarte served as the first Chair of the WTO Appellate Body and continues to serve from time to time on WTO panels and NAFTA Chapter 11 tribunals. He chaired the Uruguay Round negotiations which led to the establishment of the WTO and its dispute settlement system. This paper is based on comments made in a roundtable discussion at the Department of Foreign Affairs and International Trade, Ottawa, February 21, 2005. The views expressed are those of the author and are not to be attributed to Foreign Affairs and International Trade Canada, to the Government of Canada, or to any organization with which the author is, or has been, affiliated or associated, including the Government of Uruguay, the WTO and other international organizations.

ture, two contentious and complex agricultural subsidy cases have been brought in the WTO cotton against the United States and sugar against the European Union. Since it would not be difficult to conceive of a number of other possible challenges to US and EU agricultural subsidy practices, we can conclude that exporting countries have been, on the whole, prudent and cautious, so far. This can, in good measure, be attributed to the fact that negotiations are underway on agriculture in the Doha Round, as well as in regional agreements in which the agricultural aspects will depend on the outcome in the Doha Round. But this forbearance of litigation cannot be taken for granted and cannot be expected to continue indefinitely.

If the outcome of the Round is minimal and it does not resolve the contentious issues, or if the Round fails altogether, system friction in the WTO will grow as globalization deepens, as currency swings strain the competitive position of companies, and as global trade imbalances generate political pressures for protectionist measures. If this occurs, domestic lobbies will pressure governments to bring dispute settlement cases. In addition, domestic industries will turn to using domestic trade remedies, including anti-dumping, safeguards and countervailing duty actions. In the past few years, developing countries such as India, Mexico, South Africa and Brazil have become major users of domestic trade remedies which were once used primarily by the US, the EU, Australia, Canada and other developed countries. Domestic trade remedy actions will also likely be challenged in the WTO. These developments could result in an intolerable strain being placed on the WTO dispute settlement system.

The existing WTO Agreement is replete with issues on which negotiators could not agree and that were papered over with what is known as "constructive ambiguity" in drafting. Since the entry into force of the WTO, not one WTO Member has put forward a proposed decision on interpretation, although rules and procedures for adopting interpretations were included in the WTO Agreement precisely to help clarify the legal texts.

In many important areas, such as agriculture, subsidies and anti-dumping, in which the Uruguay Round negotiations were extremely contentious and difficult, modifications to the rules are only possible through negotiations among the Members. But if Members are not prepared to clarify the texts by adopting interpretations, or are unable to develop new rules or make amendments through successful negotiations, the burden will shift to the dispute settlement system. In this scenario, not only the caseload, but also the political stakes riding on Appellate Body decisions, would increase.

To follow this scenario a bit further, what would happen if the US Congress and the European Commission were confronted with a series of adverse Appellate Body decisions requiring changes in laws and policies to bring their measures into conformity with their respective WTO commitments? How loud and elevated would the political rhetoric become on the allegations already commonly heard, relating to the so-called "judicialization" of the trading system, the intrusion of WTO rules on domestic sovereignty, and so forth?

Moreover, given this scenario, what would happen if the major powers refused to comply with WTO Appellate Body rulings? Could the system be sustained in the face of possible non-compliance in the really difficult, political cases by the multilateral trading system's two most powerful Members? To be sure, it may be reasonably argued that this is not a perfect world, and there may be cases in which it is impossible politically for WTO Members to comply with dispute settlement rulings. In this sense, occasional non-compliance may provide the system with the flexibility it needs to accommodate the irregularities of the real world in the comparatively rigid frameworks of international agreements. However, the persistent efforts to strengthen the GATT/WTO dispute settlement system (including the 1989 measures to ensure automatic access to a panel and the measures adopted in the Uruguay Round for negative consensus to block adoption) speak to the underlying frustration that non-compliance creates for the WTO Membership.

To date, it can be said very confidently that the Appellate Body has been "unflinching" in making its determinations, regardless of what critics might say. An ambitious outcome in the Doha Round could help attenuate foreseen pressures on the multilateral trading system in a number of ways:

(a) by reducing the likelihood of non-compliance through negotiated resolution of the underlying issues;

(b) by introducing new disciplines on compliance (e.g., the suggestion that has been made by Professor William Davey to withdraw the right to request a panel as long as a Member is not itself in compliance); and

(c) by providing greater flexibility in ways to comply (e.g., compensation) or innovative approaches that give smaller WTO Members, which cannot retaliate effectively, means to re-negotiate their commitments commensurate with the degree of nullification or impairment that they have suffered from the inconsistent measure.

Finally, regional trade agreements (RTAs) will continue to proliferate unchecked if the Doha Round is not successful. This is only natural as national trade officials pursue opportunities where they may be found. But there are serious risks for the WTO and the multilateral trading system generally. The proliferation of RTAs has already undermined the WTO's foundational principle of "most favoured nation" and, in a *de facto* sense, weakened the WTO as an institution. This is because, with two minor exceptions, there have been no rulings concerning the consistency of the hundreds of RTAs notified to the WTO with Article XXIV of the GATT 1994 and Article V of the General Agreement on Trade in Services.

In the fractious environment that might follow a breakdown of multilateral negotiations, what landmines might there be for the WTO if the trade discrimination that has become so widespread as a result of the proliferation of RTAs becomes the subject of WTO disputes? Moreover, since many regional agreements include mechanisms for dispute resolution, the WTO faces a risk of "adverse selection" whereby the easy cases are

handled by the regional mechanisms while the difficult cases might be brought under the WTO Dispute Settlement Understanding (DSU). It is an axiom in legal circles that "hard cases make bad law". The ability of the WTO dispute settlement system to quietly and effectively sort out issues that have confounded negotiators would be diminished. And since the Appellate Body *must* take decisions on cases put to it (it cannot remand cases or decline to make rulings), it would find itself in the crucible created by hard cases more often than not.

There are accordingly many reasons to surmise that the Appellate Body would very strongly want a successful outcome to the Doha Round, one in which negotiators seize the thorny nettles and resolve them, rather than leaving them for the judicial organs of the WTO to address in future disputes.

A better informed understanding of the Appellate Body's role and practices

It is only human to wish to be understood. The Members of the Appellate Body are no exception to this wish. And, considered as an institution rather than as a group of individuals, the Appellate Body itself would undoubtedly benefit from a good understanding of its behaviour.

The label of "judicial activism" is sometimes applied to the Appellate Body to characterize its collective output. When used by WTO Members, this claim in often made in the context of the perspective that the Appellate Body is usurping the prerogatives of Members. When used by non-governmental organizations, it is often in the context of a claim that the Appellate Body is intruding into domestic policy space without democratic legitimacy. In point of fact, the Appellate Body has frequently exercised "judicial restraint"—ruling where it must, but refraining from ruling where it can. Under the DSU, the Appellate Body is not permitted to remand a case back to the panel for re-examination. It must approve or overturn all panel legal rulings that are appealed. Its powers allow it to modify panel rul-

ings. The Appellate Body must live with all the consequences that flow from its powers and its responsibility—including the real world impact and the criticisms that might be levelled at its reasoning. The Appellate Body has not, in fact, ruled on issues unless they were specifically appealed by a party to a dispute, and even then it has exercised judicial restraint in its rulings. It has only been as active as the Members have required it to be by bringing appeals; no more, no less.

In practice, the Appellate Body has spared no pains to maintain consensus in taking its decisions. Until recently, there had never been a dissenting opinion in any decision. Achieving internal consensus among the Members of the Appellate Body has required an enormous expenditure of time, effort and discussion to identify a way forward that permitted each and every Appellate Body Member to agree with a decision without sacrificing any principle. Given the varied backgrounds of the individuals who compose the Appellate Body (in the original group only one –that being myself—came from the ranks of trade officials), this search for consensus results in very difficult internal deliberations, particularly in complex cases such as *US - Shrimp*. It is sometimes suggested that the Appellate Body has bent to the political winds, but this is certainly not so. It is more accurate to say that the search for internal consensus in a group with disparate backgrounds has required the Members of the Appellate Body to bring to bear every possible source of argument to a case.

In the very difficult and contentious procedural area, the Appellate Body has found a reasonable accommodation in regard to "friend of the court" ("amicus curiae") submissions that aroused considerable concern among WTO Members. The Appellate Body has always kept the focus of its deliberations on the merits of the legal arguments presented by the parties and third parties in case at hand. That is as it should be—WTO disputes are matters between WTO Members.

Functional Improvements

The case load faced by the Appellate Body has far exceeded the expectations of the Uruguay Round negotiators. Moreover, the Appellate Body Members take the issues put to them very seriously and delve into them in depth, individually and through intensive and lengthy debates and discussions with their colleagues and with the supporting Secretariat staff. The time and energy required to develop a well-reasoned decision is not insignificant. There are a number of ways in which WTO Members could help ease the burden of the Appellate Body and contribute to smoother functioning of the dispute settlement system.

First, in framing arguments, it does not go over well if every possible argument is thrown in, those that have a chance of finding a favourable reading and those that do not. Even if an argument is frivolous, the Appellate Body is required by the DSU to address it, time consuming and tedious as this may be. By extension, it is of no value to one's own argument, and detrimental from the perspective of the overall process, to go to the wall in defending against each and every argument raised by the other party to the dispute.

Second, filing of massive documentation is singularly unwelcome, especially if it is of minimal relevance to the issue. Parties sometimes file voluminous evidence which turns out to be of no pertinence to the issues involved in the case. This leads to wasted effort and time. Given the Appellate Body's very short, strict time limits, this is not helpful and adds to the Appellate Body's burden in deciding the case.

Third, Appellate Body Members and Secretariat staff invest considerable time and energy in preparing for hearings. Questions that are raised in considering the issues in a given case can number in the hundreds. Particular issues that are raised can consume days of discussion and debate within the Appellate Body Division prior to the hearing. The questions for the parties that this process generates can be very penetrating—and it is not uncommon that counsel are ruthlessly pressed with difficult

questions during the hearings. An appearance before the Appellate Body can therefore be a daunting experience for the parties' counsel—even experienced lawyers of high repute have occasionally floundered. Accordingly, it is important that counsel appearing on behalf of the parties be no less well-prepared than the Members of the Appellate Body itself. It is essential to a successful hearing that counsel know the file in all its complexity by heart, be able to respond readily and confidently on any critical point, and maintain consistency and clarity in the line of argument (a muddled mind quickly reveals itself and shifting positions or lines of argument undermines credibility).

Fourth, it is much appreciated when the disputing parties act as "friends of the Appellate Body"—for example, by pointing to approaches or material that can help the Appellate Body in its reasoning.

Fifth, recognizing that cases often raise unexpected issues of systemic consequence, it is helpful to the Appellate Body when WTO Members that take an interest in the systemic issues make third party submissions—but only when they contribute independent reasoning to the issues in dispute. Third party submissions that simply support one side or the other, that reiterate well-rehearsed positions, or address issues not relevant to the case at hand make extra work for the Appellate Body and the parties without actually benefiting the system. Thus, third party participation can be, and often is, extremely useful and important, especially relating to systemic issues. But, counsel should be well-prepared and helpful to the Appellate Body and the parties when filing third party submissions.

Concluding Thoughts

In the historical evolution of the multilateral trading system, the creation of the Appellate Body represented a quantum leap in terms of deepening and entrenching the rules-based nature of the WTO. It is a fair evaluation that the Appellate Body has become a much more significant force than the negotiators in

the Uruguay Round expected, particularly in clarifying the provisions of the WTO Agreement. By the same token, the WTO Members themselves have been less forthcoming than expected in contributing to the process of clarification of ambiguous texts. Members of the WTO have, in effect, dealt with this problem by bringing issues to the Appellate Body. The stature that the Appellate Body has attained reflects this affirmation of its role by the Membership of the WTO.

It would be disingenuous to suggest that the reputation and standing attained by the Appellate Body is at all unwelcome to the individuals that comprise it. However, to operate in the milieu is to recognize the constraints and to sense the risks. What does the Appellate Body want? A fair summary would be: *first*, that the legislative arm of the WTO—the Members in their negotiating mode—pull its weight and deliver an ambitious outcome in the Doha Round, including the rules component of the negotiations; *second*, that the epistemic community convey a more nuanced appreciation and respect for the Appellate Body's responsibility and performance; and *third*, that WTO Members as parties to disputes recognize that the Appellate Body's energies are a scarce resource and approach litigation accordingly. Is that too much to ask? I think not.

A Proposal for Special Treatment in Market Access for Developing Countries in the Doha Round

Andrew Charlton[*]

Executive Summary

There is considerable dissatisfaction with the treatment given by the World Trade Organisation (WTO) to its poorest members. The development of an appropriate framework which maintains the 'rules based' trading system, but differentiates between rich and poor countries is one of the most important issues facing the Doha Round.[1]

This note outlines a proposal to increase market access for developing countries in the context of the Doha Development Agenda. It suggests that **all WTO members commit to providing free market access in all goods to all developing countries which are both poorer and smaller than themselves.** Thus all developing countries could expect free access to all countries with both a larger GDP and a larger GDP per capita.

This proposal compares favourably with alternative formulas for special and differential treatment (SDT) in market access. In particular: (i) it involves significant liberalisation; (ii) it promotes south-south liberalisation; (iii) rights and obligations are distributed progressively to all developing countries; (iv) it gives developing countries the flexibility to use trade policy to promote development by protecting their key industries from more efficient producers in larger and/or

---

[1] For recent discussions of special and differential treatment see Hoekman (2004), Keck and Low (2004), Stevens (2002), Page and Kleen (2004), and Prowse (2002).

more developed countries; (v) it complements MFN liberalisation; (vi) it brings preferential market access within the WTO, creating clear and enforceable rights; and (vii) it is simple, does not involve complex negotiations, and has built-in flexibility in the sense that national obligations change along a country's economic growth path.

The principle underlying this proposal is that all countries should participate in an enforceable system of preferential market access in which rights and obligations are distributed progressively according to objective criteria. The proposal presented in this paper represents one straightforward means of implementing this principle. Additional provisions for specific sectors, alternative dimensions to differentiate between countries, implementation periods, and various other complexities that would undoubtedly be part of any applied version of this proposal are left out of the exposition in this paper.

This proposal would require significant liberalisation by the large developed countries, particularly towards middle income countries. While this makes economic sense, it has been argued that it lies in the realm of 'wishful thinking' because of political economy constraints in the developed countries. However despite these constraints, there are good reasons to be optimistic. First, a significant share of US and EU trade is already duty free as a result of their membership of several bilateral free trade agreements. Second, there is significant support within the developed countries for free trade.[2] Third, countries would still have access to safeguard measures to protect themselves from damaging import surges.

Further, there is a lack of satisfactory alternatives within the Doha Round. Unless a proposed agreement to conclude the Round provides both significant liberalisation and significant benefits for developing countries, it will diminish the credibility of the WTO and risk being rejected by developing country Members. An SDT and market access proposal of the type presented in this paper would, on the contrary, deliver on the development promise made at Doha and involve significant liberalisation.

---

[2] See, for example, the US National Foreign Trade Council's Zero Tariff Proposal set out in Slaughter (2003).

Finally, the market access afforded by this proposal could be traded for concessions from developing countries in services, non-tariff barriers, and trade facilitation. By enriching the available trade-offs, the proposal would improve the chances for an agreement to successfully conclude the Round.

SDT in the Doha Round so Far

At the WTO's Fourth Ministerial Conference in Cancún, Mexico, in September 2003, the negotiations collapsed and the meeting ended in failure. One reason for the deadlock was that developing countries were worried about being forced into accepting obligations which would hurt their industries or impose large implementation costs on them. Many developing countries had felt disadvantaged by the last roundest of multilateral trade negotiations, the Uruguay Round, and they came to the view that no agreement was better than another bad agreement. After Cancún, developing countries stepped up their demands[3] for special and differential treatment (SDT) as a prerequisite for progress in the Round.[4]

In May of 2004, EU Trade Commissioner Pascal Lamy attempted to placate the developing countries and salvage the Round by offering a significant compromise on SDT. In a letter to trade ministers he wrote: "we propose that the least developed countries and other weak or vulnerable developing countries ... should not have to open their markets beyond their existing commitments, and should be able to benefit from increased market access offered by both devel-

---

[3] See the statement Issued by Indonesia on behalf of the G-33 member countries in Geneva, 28 July 2004.

[4] For wary developing countries, SDT was an insurance policy – it would give them the flexibility to opt out of any agreement which proved to be too onerous for them. A group of developing countries, the G33, united behind the issue of SDT. As the Round began to regain momentum in 2004, they renewed their calls for SDT to be given a higher degree of clarity and specificity. In particular, they wanted the right to identify special products of interest to developing countries on which there would be no tariff reduction commitment and no new tariff rate quota commitments.

oped and advanced developing countries. So in effect these countries should have the 'Round for Free'."[5]

The danger of the blanket approach to SDT embodied in the "Round for Free" proposal is that it creates disincentives for developing countries to participate in the Round. If the least developed countries are required to do nothing, they may be pushed to the periphery of the negotiations. Then the Development Round would bear a striking resemblance to the early rounds of trade negotiations in which the GATT operated as club for the advancement of rich country interests. In those early rounds, developing countries were burdened with few obligations, but they had only a weak voice in the negotiations and little power with which to assert their interests. Developing countries benefited from industrial country liberalization, thanks to the Most Favoured Nation (MFN)[6] principle, but their peripheral role meant that they could exert little pressure on the *way* that industrialised countries liberalized. Thus liberalization of goods of interest to developed countries proceeded swiftly, but goods of interest to developing countries, especially labour-intensive goods, lagged behind and developing countries ultimately suffered. Some developed countries were happy with this system because the small and poor countries did not have sufficiently attractive markets to bother with: the benefits of market access were smaller than the costs of liberalizing their own labour-intensive import-competing sectors. The "Round for Free" approach smacks of the same two-tiered

---

[5] See the letter of 9 May 2004 by Pascal Lamy and EU Agriculture Commissioner Franz Fischler. Lamy quickly stepped back from this offer. In June he noted that the "Round for Free" slogan was perhaps a misnomer since developing countries would be required to make commitments on binding their tariffs in some areas, and participating in negotiations in trade facilitation. He coined the somewhat less catchy slogan "Round at a modest price". See the speech by Pascal Lamy, European Commissioner Responsible for Trade: "Where next for EU Trade policy?", Deutsche Gesellschaft für Auswärtige Politik, Berlin, 11 June 2004.

[6] MFN is enshrined in the first article of the General Agreement on Tariffs and Trade (GATT), which governs trade in goods. MFN is also a priority in the General Agreement on Trade in Services (GATS) (Article 2) and the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS) (Article 4), although in each agreement the principle is handled slightly differently.

system which exempted developing countries from commitments but excluded them from the negotiations. However, as Keck and Low (2004) argue, "where new policy areas or new rules are under negotiation, or consideration for negotiation, the best interests of developing countries would be served through engagement with respect to the substance of core proposals." Another problem with the "Round for Free Approach," which concerns many negotiators from developed countries, is that it allows the poorest countries to continue to participate in the Round (and indeed to slow its progress, as per one interpretation of the failure at Cancún) when they are contributing very little to it.

Moreover, the "Round for Free" approach may result in substantial opportunity costs for developing countries by robbing them of the benefits of liberalisation of south-south trade. Developing countries now account for around one-third of global trade. Intra-developing country, or 'south-south', merchandise trade has grown at twice the pace of world trade over the past decade. Yet barriers to south-south trade are high.[7] Indeed, 70 percent of the tariffs paid by developing countries (US$57 billion annually) are paid to other developing countries. Thus, developing countries stand to benefit a great deal from improved market access to other developing countries. The World Bank (2002) estimates that developing countries stand to realise welfare gains of more than US$30 billion per year if other developing countries eliminated tariffs on industrial goods and a further US$30 billion if they removed their barriers to agricultural trade.

A Doha Round Market Access Proposal (Doha Map)

Thus the challenge is to design special and differential treatment which gives developing countries flexibility to deal with their development problems and minimises adjustment and implementation costs, without marginalising their participation in the global trading system or foregoing the gains from south-south liberalisation. To achieve this, all WTO members could commit to providing free market access in all goods to

---

[7] For example, East Asian exporters face tariffs in other East Asian countries that are 60 percent higher than in rich nations.

all developing countries poorer and smaller than themselves. Thus all developing countries could expect free access to all markets with (a) a larger GDP and (b) a larger GDP per capita. This special and differential treatment provision would bind developing and developed countries alike. For example, a middle income country like Egypt with GDP per capita of US$1,390 and GDP of US$82 billion, would receive free market access to countries like the United States, but would be required to give free market access to a country like Uganda (GDP per capita of US$240 and GDP of US$6.2 billion).

This proposal has several advantages over alternative schemes:

It involves significant liberalisation

The quantity of liberalisation delivered by Doha MAP is in part a function of the correlation coefficient of the GDP and GDP per capita of WTO Members—in the limit case where the coefficient is -1, MAP requires no liberalisation from any country. Figure 1 plots the GDP and GDP per capita of WTO Members. The correlation is large and positive (0.6), implying that MAP delivers significant liberalisation. The dotted lines in Figure 1 illustrate the implications of the proposal for Egypt, a country in the middle of the distribution of both size and wealth. Doha MAP would require Egypt, after a negotiated implementation period, to provide free market access to more than 50 developing countries to its south-west in Figure 1 (with total market size of US$500 billion). In return it would receive free market access to more than 20 developed and upper-middle-income countries to its north-east in Figure 1 (with a total market size of US$28 trillion).[8]

As with existing preferential schemes, the effect of Doha MAP would be, to some extent, limited by rules of origin. It is beyond the scope of this note to describe options for rules of origin that might be required to implement Doha MAP, but it is worth noting that MAP would significantly reduce the distortionary effect of rules of origin on LDCs' trade

---

[8] In this calculation the EC is treated as one Member; this makes little difference to the market size numbers.

compared to current SDT approaches since many of the middle income countries from which LDCs might import intermediate inputs would also receive preferential access under MAP to some of the rich countries to which the final goods are exported. Thus, while not eliminating the problem of rules of origin in SDT, MAP would reduce its effect in practice compared to the status quo.

Figure 1: WTO Members' GDP and GDP per capita
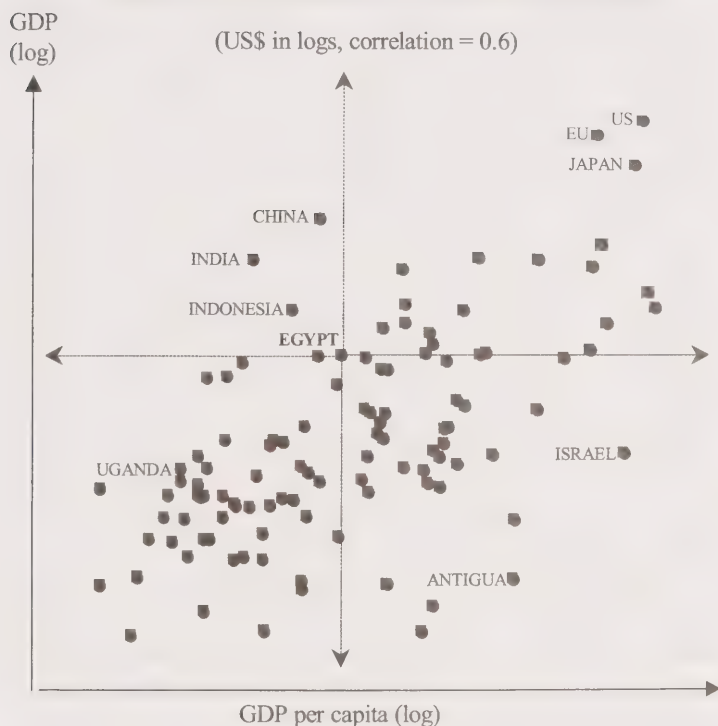
Particularly south-south liberalisation

Another advantage of the MAP proposal is that is takes south-south liberalisation seriously. Many existing types of SDT (including several proposed changes) do little to promote south-south trade, the liberalisation of which could bring large gains to developing countries. Indeed most Doha Round estimates indicate that the scope for welfare gains for developing countries is larger from the liberalisation by other de-

veloping countries than from liberalisation by developed countries (Francois, van Meijl, and van Tongeren 2004).

South-south liberalization has progressed slowly. Attempts at preferential market access agreements have been made outside the WTO under the auspices of the Global System of Trade Preferences among Developing Countries (GSTP).[9] Unfortunately, the GTSP is based on reciprocity—one reason for the low participation of least developed countries among its members —and it has struggled to make significant progress.[10] Bilateral and regional free trade agreements (FTAs) between developing countries are increasing in number, but it should not be assumed that south-south FTAs are unequivocally good for development since they discriminate against third party developing countries and the margin of discrimination is higher than is the case in north-north FTAs because developing country MFN tariffs tend to be higher. Thus, there may be a strong case for introducing a development dimension into south-south agreements. There are schemes being considered by some larger developing countries including India, China, and Brazil which would give special access to the least developed countries. While additional market access would be welcome, these schemes, like the existing GSP schemes operated by the advanced industrial countries, would be a patchwork of discretionary and conditional promises rather than clear legal rights enforceable within the WTO.

Within the WTO, developing countries have often been urged to reduce their MFN tariffs on the grounds that this would lead to an increase in south-south trade. Doha MAP recognises that, for this purpose, liberalisation need not occur on an MFN basis.

---

[9] The GSTP, established in 1988 and promoted by UNCTAD, provides trade preferences to developing countries without extending them to developed countries.

[10] The two previous GTSP rounds, in the past two decades, were not as successful as expected, due to the economic situation of the poorest developing countries and the poor negotiating capacity of member states.

Obligations are distributed progressively

Doha MAP is progressive in the sense that it requires significant south-south liberalisation from middle income countries and very little from the poorest and most vulnerable countries. It requires the most liberalisation from the countries in the North-East of Figure 1 (in particular the Quad countries) and less of those in the South-West of the Figure 1 (mostly African LDCs). Under this scheme, all but the very poorest countries do *not* get the "Round for free" since all countries accept the obligation to provide market access to other WTO Members smaller and poorer than themselves. In return the developing countries receive considerably more market access, under well defined commitments, than under existing preferential schemes which are discretionary schemes operated by industrialised countries that are not subject to detailed WTO regulation governing their implementation.

Figure 2 plots the ratio of market access rights to market access obligations under MAP for developing country WTO Members against their GDP per capita.[11] The vertical axis is the ratio of the sum of the size of all the (bigger and richer) markets to which the country will have free access to the sum of the size of all the (poorer and smaller) markets to which the country will be obliged to give free market access. The ratio of rights to obligations is progressive in the sense that the poorest countries get free access to huge new markets while being obligated to give free access to much smaller markets. In figure 2, the extent of progressivity of the proposal is reflected in the extent to which the scatter tends to slope down from left to right with a small, poor country like Sierra Leone having amongst the highest ratios and larger, richer economies like Mexico having among the lowest ratios.

As Table 1 indicates, all developing country WTO Members benefit from the scheme. Even the largest and

---

[11] Figure 2 describes free market access rights and obligations after the implementation of MAP, rather than the change in rights and obligations, i.e., it includes preferential access under existing schemes, as well as new access.

richest developing countries[12] receive free access to markets whose total size is more than seven times the size of the markets to which they must give free access. The median ratio of market access rights to obligations under MAP is 303, i.e., the median developing country receives access to markets 303 times the size of the markets to which it must *give* free access. Alternatively, measured by imports, the median developing country receives free market access to countries whose total imports are 113 times the size of the imports of the countries to which it is required to give access.

Figure 2: A progressive distribution of rights and obligations

(Ratio of new market access rights to obligations for developing countries)

GDP per capita (log)

Countries can manage major import threats

The proposal imposes no extra obligation on developing countries to open their markets to larger or more developed economies. This gives developing countries the *option* to

---

[12] For the purposes here, developing countries are defined as those that had a GDP per capita below US$10,000 in 2003.

provide their key industries with some protection from imports from economies with cost advantages derived from either scale of their economies (e.g. larger countries, particularly China), or technological advantages (more developed countries). The prevalence of import barriers in some sectors indicates that developing countries wish to protect particular industries from import competition. This protection may form part of an industrial strategy based on 'infant industry' protection[13], but is more likely to originate in a desire to avoid adjustment costs which could be particularly severe in developing countries characterised by high unemployment, weak risk markets, and low levels of social insurance.

As Hoekman et al. (2003) point out, protectionism can be self defeating for developing countries in a world where multinational corporations have made production increasingly fragmented internationally. For many developing countries, "the only option to reach the minimum scale required for sustained growth in output is integration with the rest of the world" (Keck and Low 2004).

This proposal facilitates integration by providing all developing countries with significantly increased market access to larger and richer markets, while providing the option of protection from imports from countries which are at later stages of development or have scale advantages.

Is consistent with other MFN liberalisation schemes

It is important to point out that this proposal is not anti-openness. It does not involve any increases in existing MFN rates. Each country would continue to uniformly apply MFN rates to larger and more developed countries.

In addition, this proposal is squarely in the realm of SDT. There is still a role for the WTO to negotiate MFN tariffs, i.e., it complements other proposed modalities for MFN tariff reduction, rather than replacing them.

---

[13] Although this would not be particularly desirable from a theoretical point of view and, indeed, seems unlikely to be a major motivating factor since infant industry policies have been shown to be ineffective (Baldwin 1969) and inferior to alternative forms of industry assistance (McCulloch, Winters and Cirera 2002).

One concern with the proposal is that it may affect the bargaining positions of developing countries in future rounds. One of the unfortunate side-effects of existing preferential schemes is that they create an inbuilt incentive for developing countries to block MFN liberalisation which would erode their preference margins. But this problem is much less severe for MAP than for existing preference schemes because MAP is far less distortionary: large rich countries do not give LDCs preferences that they do not also give to middle income countries. Thus MFN liberalisation by developed countries does not cause LDCs to lose out relative to middle income countries.

It transforms discretionary preferential schemes into well defined obligations within the WTO

One of the main advantages of this proposal over existing types of SDT is that it delivers clearly defined and legally binding rights to developing countries in a way that existing preferences do not. Many of these existing preference schemes were originally spawned by Part IV of the GATT which includes provisions on preferential treatment for developing countries. This exception was further expanded in 1979 in the decision which has come to be known as the 'Enabling Clause'. This consolidated the concept of "differential and more favourable treatment" for developing countries as well as the principle of non-reciprocity in trade negotiations.[14]

However, the problem with this (potentially) wide-ranging clause is that is has never placed any formal obligations on developed countries. Instead, piecemeal preferential deals have been established which cover a limited range of goods from a limited group of countries. These preferences, the most important of which are offered by the "Quad coun-

---

[14] The most significant provision of the Enabling Clause is that which enables Members to accord differential and more favourable treatment to developing countries as a departure from the MFN Clause. It stipulates that "contracting parties may accord differential and more favourable treatment to developing countries, without according such treatment to other contracting parties".

tries" (US, EU, Canada, Japan) often divert trade from other poor countries.

Another problem with preferential schemes is their uncertainty. Keck and Low (2004) argue that SDT should enshrine "legal rights and obligations", whereas existing preferences have become merely "legally unenforceable statements of intent or best-endeavour undertakings." Preferences are not binding on the countries which grant them, and can be altered to exclude certain products or withdrawn entirely at the discretion of the preference-provider. For example, in 1992 the US withdrew $60 million worth of pharmaceutical imports from their preference scheme because the US Trade Representative determined that India had weak patent protection which adversely affected US companies.

Without binding obligations, preference-providers have faced pressure from their own import-competing domestic lobbies to minimise the scope of their preferential schemes. As the Sutherland (2005) report notes, it is "grantor, rather than grantee, country interests [which] have determined the product coverage and the preference margins in GSP schemes."

Balances simplicity against the need to differentiate

The idea that SDT should be provided to countries based on objective access criteria has been previously addressed by Stevens (2002). Stevens' proposal suggests that a new SDT regime involve "greater differentiation of treatment between WTO members which, in turn, implies the establishment of objective criteria on which to determine the differentiation." For example, he suggests that access to some types of SDT in agriculture should be based on measurable criteria relating to food security, i.e., countries could qualify for special treatment if they have per capita caloric intake of less than a certain level (indicating vulnerability to food insecurity) and a high share of agriculture in GDP (indicating the importance of agriculture in livelihoods) and a high share of food imports to GDP (indicating import dependency). Stevens' approach involves setting objective criteria on an agreement-by-agreement basis. Thus, special and differential treatment would be available to countries which met objective precon-

ditions indicating their need for exception and/or assistance. The appeal of this approach is that it closely matches the needs of specific countries to special treatment in different provisions. In addition, it provides more certainly to developing countries since, once the conditions are predetermined; eligibility would be automatic rather than at the discretion of other WTO Members.

The disadvantage of this approach is that it would add to the complexity of trade negotiations and greatly increase transactions costs. SDT measures are already overly complicated in many areas. Hudec (1987) refers to preferences as systems of "refined complexity", determined by an "orgy of fine-tuning". The process of tailoring objective criteria for SDT in each agreement requires countries to agree on measurable criteria and agree on eligibility cut-offs. As Stevens (2002) himself notes, "the whole process is likely to be fraught with political difficulty". It is likely that neither the international consensus on these issues nor the necessary negotiating capacity currently exists to operate such an ambitious and resource-intensive SDT system.

By contrast, the MAP proposal is simple to negotiate. It would entirely do away with the whole 'spaghetti-bowl' of GSP preferences (although not FTA preferences) and it would save the EC the bother of negotiating the market access part of the economic partnership agreements with African, Caribbean and Pacific States.[15]

Moreover it includes an inbuilt flexibility that removes the need for renegotiation over time. As countries develop and overtake others, they will, after an implementation period, lose some preference rights and accept obligations to poorer countries. Alternatively, the scheme could be designed to include a 'one-way' provision so that free trade would be monotonically increase in a dynamic world where rankings change.

---

[15] The Cotonou Partnership Agreement between the African, Caribbean and Pacific States (ACP) and the European Union, signed June 2000, mandates the negotiation by 2007 of a series of economic partnership agreements (EPAs) between regional groups of ACP States on the one hand and the EU on the other.

Concluding Remarks

Developing countries have been understandably reluctant to commit to large reductions in their tariff levels. They are concerned that open borders will lead to a flood of cheap imports from more efficient producers which could destroy their fledgling industries before they have a chance to develop. Because they are already characterised by high unemployment and weak private and social insurance, many developing countries believe that the adjustment costs from significant MFN tariff reduction are too large to be seriously considered. Consequently developing countries have not offered large reductions in border protection in WTO negotiations. As a result, south-south trade has suffered, and developing countries have little bargaining power in their negotiations with developed countries.

The proposal in this paper distributes new market access progressively, ensuring that the largest gains accrue to the smallest and poorest countries, and it distributes liberalisation obligations progressively, requiring that the largest and richest countries liberalise most. There are many other issues associated with tariff reduction in poor countries including adjustment costs and declining revenue to governments. These issues, and their implications for the need for technical and development assistance to poor countries, are not discussed here. The advantage of the market access proposal outlined in this note is that it provides significant liberalisation, does not demand reciprocity from poor nations to richer ones, and places simple and well-defined obligations on both rich and poor countries alike.

References

Baldwin, R. E. (1969). "The Case against Infant-Industry Tariff Protection," *Journal of Political Economy*, 77, no. 3, pp. 295-305.

Francois, J., H. van Meijl, and F. van Tongeren, (2004). "Trade liberalisation in the Doha Round" *mimeo*, Erasmus University.

Hoekman, B. (2004). "Operationalizing the Concept of Policy Space in the WTO", *World Bank Policy Brief* No. 4, July 2004.

Hoekman, B., Michalopoulos, C., Winters, L. A. (2003). "Special and Differential Treatment for Developing Countries: Towards a New Approach in the WTO", *mimeograph,* World Bank.

Hudec, R. E. (1987). *Developing Countries in the GATT Legal System*, Thames Essays, London: Trade Policy Research Centre.

Keck, A., and Low, P. (2004). "Special and Differential Treatment in the WTO: Why, When and How?" *Staff Working Paper* ERSD-2004-03, Economic Research and Statistics Division World Trade Organization, Geneva.

McCulloch, N., L. Winters and X Cirera (2002). *Trade Liberalisation and Poverty: A handbook*. London, Centre for Economic Policy Research.

Page, S. and P. Kleen. (2004). *Special and Differential Treatment of Developing Countries in the World Trade Organization*. Overseas Development Institute, Paper prepared for the Ministry of Foreign Affairs, Sweden.

Prowse, S. (2002). "The Role of International and National Agencies in Trade-related Capacity Building", *The World Economy* 25, 9: 1235-1261.

Slaughter, M. J. (2003). "Tariff Elimination for Industrial Goods: Why the Gains Will Far Outweigh Any Losses", Background Paper prepared for the US National Foreign Trade Council, http://www.wto.org/english/forums e/ngo_e/posp44_nft c_tarif_paper_e.pdf.

Stevens. C. (2002). "The future of Special and Differential Treatment (SDT) for developing countries in the WTO", Institute of Development Studies (IDS) Working Paper 163, Brighton, Sussex: IDS.

Sutherland, P. (2005). *The Future of the WTO. Addressing institutional challenges in the new millennium,* Report by the Consultative Board to the Director-General Supachai Panitchpakdi. World Trade Organisation, Geneva.

World Bank, (2002). *Global Economic Prospects and the Developing Countries: Making Trade Work for the Poor.* World Bank, Washington.

World Bank, (2004). *Global Economic Perspectives 2004: Realizing the Development Promise of the Doha Agenda.* World Bank, Washington.

Table 1: Market Access Rights and Obligations for Developing countries*
under MAP

Country	Code	GDP (US$ bn)	GDP per capita (US$)	Ratio of rights and obligations	
				ratio of GDP^	ratio of imports#
Mexico	MEX	626	6,230	7	2
Poland	POL	210	5,270	13	4
Turkey	TUR	238	2,790	18	6
China	CHN	1,410	1,100	19	7
Czech Rep	CZE	85	6,740	19	6
Brazil	BRA	492	2,710	20	6
Hungary	HUN	83	6,330	22	7
South Africa	ZAF	160	2,780	25	8
Argentina	ARG	130	3,650	26	9
Malaysia	MYS	103	3,780	26	9
Venezuela	VEN	85	3,490	29	10
Thailand	THA	143	2,190	33	12
Chile	CHL	72	4,390	35	12
Egypt	EGY	82	1,390	56	21
Romania	ROM	60	2,310	56	20
Slovenia	SVN	26	11,830	58	15
Croatia	HRV	28	5,350	58	16
Slovak Rep.	SVK	32	4,920	60	18
Peru	PER	61	2,150	61	23
Colombia	COL	78	1,810	62	26
Philippines	PHL	81	1,080	80	37
Lithuania	LTU	18	4,490	84	26
Indonesia	IDN	208	810	86	38
Costa Rica	CRI	18	4,280	89	29
India	IND	599	530	97	42
Tunisia	TUN	24	2,240	101	32
Bulgaria	BGR	20	2,130	116	40
Panama	PAN	13	4,250	123	40
Guatemala	GTM	25	1,910	123	44
Trinidad	TTO	10	7,260	126	32
Ecuador	ECU	27	1,790	132	49
El Salvador	SLV	14	2,200	133	47
Dominican R.	DOM	16	2,070	134	49
Morocco	MAR	45	1,320	137	51
Uruguay	URY	11	3,790	149	47
Pakistan	PAK	69	470	152	78
Estonia	EST	8	4,960	161	51
Latvia	LVA	10	4,070	167	54
Sri Lanka	LKA	19	930	173	74

Jamaica	JAM	8	2,760	209	69
Jordan	JOR	10	1,850	211	74
Bangladesh	BGD	52	400	217	103
Botswana	BWA	7	3,430	218	72
Angola	AGO	13	740	251	108
Cote d'Ivoire	CIV	14	660	260	113
Bolivia	BOL	8	890	294	117
Cameroon	CMR	12	640	301	125
Albania	ALB	6	1,740	303	107
Honduras	HND	7	970	304	120
Gabon	GAB	6	3,580	307	101
Mauritius	MUS	5	4,090	345	105
Paraguay	PRY	6	1,100	354	127
Kenya	KEN	14	390	394	176
Namibia	NAM	5	1,870	396	132
Senegal	SEN	7	550	411	186
Nigeria	NGA	50	320	491	241
Ghana	GHA	8	320	589	292
Georgia	GEO	4	830	812	298
Nicaragua	NIC	4	730	818	300
Tanzania	TZA	10	290	947	457
Congo, Rep.	COG	4	640	1,008	351
Armenia	ARM	3	950	1,218	375
Barbados	BRB	3	9,270	1,239	252
PNG	PNG	3	510	1,540	582
Zambia	ZMB	4	380	1,615	783
Benin	BEN	4	440	1,638	659
Guinea	GIN	4	430	1,739	713
Uganda	UGA	6	240	1,799	833
Nepal	NPL	6	240	1,800	834
Swaziland	SWZ	2	1,350	1,922	534
Madagascar	MDG	5	290	1,997	938
Mali	MLI	4	290	1,998	939
Cambodia	KHM	4	310	2,015	1,091
Haiti	HTI	3	380	2,051	910
Moldova	MDA	2	590	2,495	908
Burkina Faso	BFA	4	300	2,739	1,349
Kyrgyz Rep.	KGZ	2	330	4,791	2,232
Togo	TGO	2	310	5,096	2,470
Mozambique	MOZ	4	210	5,186	2,232
Antigua	ATG	1	9,160	6,098	1,052
Chad	TCD	3	250	6,306	2,626
Niger	NER	3	200	9,328	3,110
Lesotho	LSO	1	590	9,683	3,207
Mongolia	MNG	1	480	9,873	3,267

Maldives	MDV	1	2,300	13,242	4,067
Mauritania	MRT	1	430	15,263	5,307
Central Af Rep	CAF	1	260	18,829	6,894
Rwanda	RWA	2	220	18,837	6,900
Malawi	MWI	2	170	18,841	6,903
Guyana	GUY	1	900	19,933	7,798
Grenada	GRD	0	3,790	24,227	6,320
Djibouti	DJI	1	910	35,104	11,292
Sierra Leone	SLE	1	150	35,360	15,037
Dominica	DMA	0	3,360	117,865	26,270
Solomon Is.	SLB	0	600	131,886	33,814
Gambia.	GMB	0	310	135,371	34,859
Burundi	BDI	1	100 na	na	
Guinea-Bissau	GNB	0	140 na	na	
Congo, DR	ZAR	6	100 na	na	

* Developing countries are defined as all countries with GDP per capita of less than US 10,000 in 2003.

^ Ratio of market size of countries to which free market access would be granted to market size of countries to which this country would be required to give free market access.

Ratio of imports of countries to which free market access would be granted to imports of countries to which this country would be required to give free market access.

Part II

Staff Papers and Commentaries

The Determinants of Canadian Direct Investment Abroad

Eugene Beaulieu, Shenjie Chen
and Erica Pohjola[*]

Abstract

It is well known that, based on any measure of economic openness, Canada is a very open economy. However, until recently, governments and analysts have all but ignored a large and important aspect of Canadian economic activity: Canadian investment in other countries. This Chapter provides a comprehensive analysis of Canadian Foreign Direct Investment Abroad (CDIA) with a focus on empirically examining the key factors that determine CDIA. The trends and patterns of CDIA from 1987 to 2003 are examined in light of a survey of the literature on the determinants of foreign direct investment. We then provide an overview and analysis of an extensive dataset that we have constructed of variables used to measure "host country" attributes. We develop an econometric framework based on the literature review and the gravity model to empirically examine the determinants of CDIA. The econometric analysis finds that the key characteristics of host countries that positively impact CDIA are geographic proximity to Canada, size of the economy, and the extent of investor protection.

[*] Eugene Beaulieu is an Associate Professor in the Department of Economics at the University of Calgary and was the Norman Robertson Fellow and Visiting Scholar at International Trade Canada over the academic year 2004-5. He was also visiting Carleton University during that period. Contact: beaulieu@ucalgary.ca.

Shenjie Chen is a Senior Policy Research Analyst in the Office of the Chief Economist (CSE), Foreign Affairs and International Trade Canada. Contact: Shenjie.Chen@international.gc.ca;

Erica Pohjola is a Policy Research Analyst in the Office of the Chief Economist (CSE), Foreign Affairs and International Trade Canada. Contact: erica.pohjola@international.gc.ca.

Introduction

Based on any measure of economic openness Canada qualifies as a very open economy. Canada exports approximately $400 billion[1] worth of goods, and imports over $350 billion each year, which is equivalent to between 30 and 40 percent of GDP. The flow of immigrants into Canada is over 200,000 per year, and the stock of foreign direct investment in Canada is over $350 billion. However, until recently, governments and analysts have all but ignored a large and important aspect of Canadian international activity: Canadian outward investment. In 2003, the total stock of Canadian direct investment abroad (CDIA) amounted to just over $403 billion and spanned over 100 countries around the world. During the same year, the stock of foreign direct investment (FDI) in Canada was just over $364 billion. Between 1987 and 2003, CDIA grew almost 400 percent from $74 billion in 1987. Although the share of CDIA in poor countries doubled from around 12 percent in 1987-91 to over 24 percent in 1999-03, the wealthy industrial countries such as the U.S., the EU and Japan are host to the lion's share of CDIA.

The primary focus of Canadian international commercial policy has historically been on trade (exports and imports) and on inward direct investment. But recently, more attention has been paid to Canada's direct investment abroad. The 2005 Canadian International Policy Statement argued that outward investment provides a number of benefits to Canadians: increased competitiveness, R&D, technology transfers, and positive externalities. The Statement pointed out that, to support such investments, businesses must be assured that they will receive predictable and equitable treatment in the host country. The Department of International Trade's website notes that emerging markets involve both opportunities and challenges; it poses the question: "Is enough being done to encourage outward as well as inward investment?"

A prerequisite to answering this question is to understand the patterns and determinants of CDIA. Unfortunately, the de-

---

[1] Unless otherwise noted, data are reported in Canadian dollars.

terminants of investment—particularly outward investment—are still not well understood.

Neoclassical economic theory predicts that capital should flow from capital-abundant to capital-scarce countries. Under the standard assumptions of neoclassical economics, such as perfectly mobile capital, the return to investment in every country should be identical. Lucas (1990) compares the US and India in 1988 and argues that according to the neoclassical model, all capital should have flowed to India, as the marginal product of capital in India was roughly 58 times that of the US. However, this pattern of direct investment was not observed between the US and India, nor do we observe it in the CDIA data.

Based on an overwhelming amount of quantitative evidence that is inconsistent with the theory, Lucas questions the neoclassical model's assumptions. But if this is the case, what is amiss with the model? Why doesn't capital flow from a rich country like Canada to poor countries? What then determines the pattern of Canadian direct investment abroad?

The main theoretical explanations for this "Lucas paradox" can be grouped into two categories. The first group of explanations focuses on market failures in recipient countries — mainly sovereign investment risk and asymmetric information. The argument is that although capital has a high return in poor countries, it does not flow there because of market failures. Gertler and Rogoff (1990) show that asymmetric information problems can cause a reversal in the direction of capital flows relative to the perfect information case. If investors lack good information about a potential market, they tend to invest less in that market relative to what the neoclassical model predicts, as the model does not take into account the information asymmetry.

Similarly, sovereign risk is not included in the neoclassical model, but investors do include it in their consideration of whether or not to invest in a country. Countries with better institutions, such as secure property rights and honest, stable governments, generally have higher levels of investment in physical and human capital. The literature on the determinants of FDI finds that FDI is positively associated with government size,

political stability and openness. Wei and Wu (2002) find that, *ceteris paribus*, corrupt countries receive substantially less FDI.

The second group of explanations focuses on "missing factors" that affect the marginal product of capital via the production function. These explanations account for the lack of capital flows from rich to poor countries by looking at the existence of other factors that positively affect the returns to capital but are generally ignored by conventional neoclassical economic theory. For example, if human capital positively affects the return to capital, then less capital will flow to countries with lower endowments of human capital. In his work, Lucas considers the influence of both differences in human capital quality and the role of human capital externalities. He finds that accounting for the differences in human capital quality significantly reduces the differentials in the estimated returns to capital; including the role of externalities eliminates those differentials altogether.

We extend Lucas's "missing factors" approach to include market size, resources, and government polices. We construct a panel database covering 40 countries from 1987 to 2003. We use these data within a gravity-model framework to examine the economic, institutional, political and policy-based factors that determine the pattern of Canadian FDI across countries and over time. We find that the key host country characteristics that impact CDIA are: geographic proximity to Canada; size and growth of the economy; and taxation policy. We find that having an investment agreement with a country does not have a statistically significant effect on CDIA.

The remainder of the paper is structured as follows. The next section analyzes recent trends and patterns in CDIA from 1987 to 2003. The third section presents a literature review on the determinants of FDI in general. The literature review is structured with Lucas's (1990) question regarding where FDI locates, in mind. We develop a framework based on the literature review and the gravity-model of international transactions. This provides a framework for the empirical analysis that follows. The empirical analysis is presented in the fourth section. The fifth section discusses the results and provides some concluding remarks.

Trends and patterns in CDIA

Table 1 presents the distribution of CDIA across countries over two periods (averaged over the 1987-91 and 1999-03 periods). The top destinations for CDIA were: the United States, the United Kingdom, countries in Non-US North America (mainly Bahamas, Barbados, and Bermuda); Ireland, the Netherlands, and Hungary in Europe; Japan in Asia; and Brazil and Chile in South America. In Europe, CDIA was concentrated in the Netherlands and two low-tax countries (Ireland and Hungary), rather than in big continental European countries such as Germany and France. CDIA in Central and South America was substantial, particularly in resource-rich Brazil and Chile. CDIA in each of these countries was larger than that in Mexico. In non-Japan Asia, CDIA focused on resource-rich Indonesia and two Asian financial centers: Singapore and Hong Kong. CDIA in two major Asian emerging markets, China and India, was relatively small. Average CDIA in China during the period of 1999-2003 was only $602 million, which was smaller than Canadian investment in Colombia or New Zealand. CDIA in India over the same period was $178 million, roughly the amount of CDIA in Panama over the same period.

Canadian direct investment abroad has more than tripled since the late 1980s to reach $429.6 billion by the end of 2002. Over the same period, the geographic pattern of CDIA changed substantially. The United States remained the most favored location for CDIA, accounting for 47.1 percent of CDIA during the period of 1999-2003, though its share in total CDIA fell gradually from the high of 62 percent over 1987-91. The share of non-US North America in total CDIA more than doubled, increasing from 6.7 percent in 1987-91 to 14.2 percent in 1999-2003, and that of Europe rose from 20.7 percent to 24.7 percent over the same periods. Despite the extensive media coverage on the surge in FDI flowing to East and Southeast Asian economies, Canadian direct investment in that region was flat. Canadians made more investment in South and Central America than in East and Southeast Asia.

Table 1: Top Destinations for Canadian Direct Investment Abroad

	Ave (1987-91)	Ave (1999-03)	Percent of Total
United States	55,981	176,450	47.1
United Kingdom	11,205	36,307	9.7
Barbados	1,351	23,136	6.2
Ireland	1,173	11,763	3.1
Netherlands	1,188	10,316	2.8
Bermuda	1,769	9,823	2.6
Bahamas	1,872	7,738	2.1
Japan	890	7,023	1.9
Cayman Islands	143	6,827	1.8
Hungary	2	6,807	1.8
Brazil	1,521	6,557	1.7
Germany	800	5,917	1.6
Chile	230	5,704	1.5
France	1,456	5,634	1.5
Australia	2,091	5,131	1.4
Argentina	128	4,924	1.3
Switzerland	978	4,039	1.1
Singapore	1,795	3,730	1.0
Indonesia	997	3,668	1.0
Mexico	218	3,293	0.9
Hong Kong	522	3,134	0.8
Italy	380	2,705	0.7
Belgium	535	2,651	0.7
Peru	9	1,924	0.5
Sweden	11	1,268	0.3
Spain	472	959	0.3
Thailand	33	918	0.2
South Korea	28	821	0.2
Luxembourg	9	800	0.2
Colombia	25	794	0.2
New Zealand	152	675	0.2
Austria	26	645	0.2
China	16	602	0.2
Jamaica	300	582	0.2
Malaysia	81	560	0.1

112

In considering the "Lucas paradox" in the Canadian context, it is important to recognize that the direction of CDIA has shifted in the past several decades. In the early 1950s, developing countries accounted for 10 percent of CDIA. That share reached 23 percent by 1970. The 1980s was a lost decade for developing countries; their share of CDIA slipped back to the range of 13-15 percent. Since 1989, however, the developing country share of CDIA surpassed the previous peak in 1970 to reach 24.4 percent in 1999. So the Lucas paradox in the Canadian context is one of degree: why doesn't more CDIA go to developing countries?

Table 2 presents the geographical distribution of CDIA among developing countries. The distribution of CDIA was very uneven. Barbados, Bermuda, Bahamas, and the Cayman Islands top the list, followed by Hungary and several resource-rich South and Central American economies including Brazil, Chile, and Argentina. Canadian investment in developing Asia was concentrated in Indonesia and two financial centres, Singapore and Hong Kong. These countries plus Mexico represented 90 percent of CDIA in developing countries.

A surprising development is the dramatic rise of CDIA in Hungary, which increased from an average of $2 million in 1987-91 to $6.8 billion in 1999-2003 (see Table 3). CDIA in Peru, the Cayman Islands, China, Argentina, Costa Rica, and Colombia also exhibited very strong growth. Some advanced industrialized countries, including Sweden and Luxembourg, also witnessed large inflows.

By far the largest share of CDIA was concentrated in finance and insurance, followed by energy and minerals, services, and machinery. CDIA was more diversified in industrial countries than in developing countries (see Figures 2 and 3). 50 percent of CDIA into industrial economies was in services, 24 percent in "other" sectors and 19 percent in energy and minerals, with 7 percent in machinery and transportation equipment. CDIA in developing countries, on the other hand, was more concentrated, with services making up 60 percent of CDIA, energy and metallic minerals 27 percent, while machinery and transport equipment (representing most of manufacturing activities) accounted for only 2 percent.
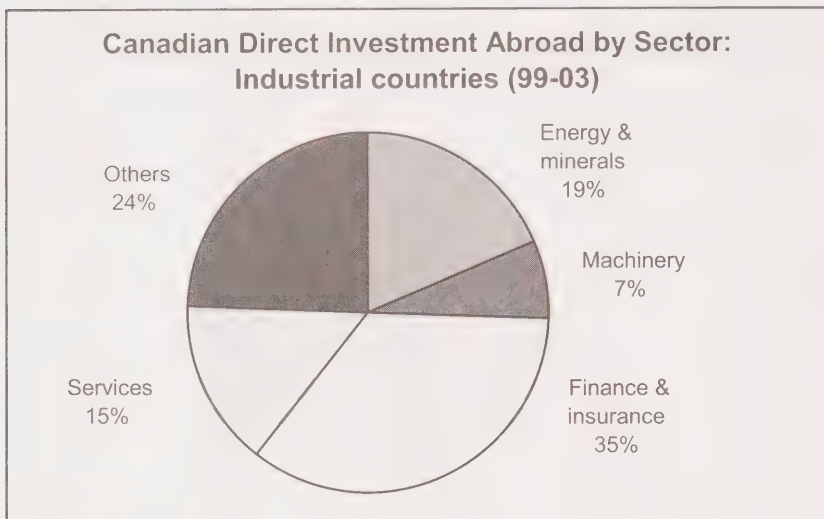
Table 2: CDIA in Developing Countries

	Ave (1987-91)	Ave (1999-03)	Growth (%)
Barbados	1,351	23,136	1,613
Bermuda	1,769	9,823	455
Bahamas	1,872	7,738	313
Cayman Islands	143	6,827	4,674
Hungary	2	6,807	340,240
Brazil	1,521	6,557	331
Chile	230	5,704	2,378
Argentina	128	4,924	3,747
Singapore	1,795	3,730	108
Indonesia	997	3,668	268
Mexico	218	3,293	1,413
Hong Kong	522	3,134	500
Peru	9	1,924	21,282
Thailand	33	918	2,724
South Korea	28	821	2,868
Colombia	25	794	3,050
China	16	602	3,781
Jamaica	300	582	94
Malaysia	81	560	595
Turkey	NA	549	NA
Papua New Guinea	86	269	212
Russian Federation	NA	265	NA
Philippines	30	258	764
Venezuela	81	247	205
Kazakhstan	NA	230	NA
Ecuador	9	212	2,253
Taiwan	136	197	44
Panama	17	187	974
India	78	178	129
South Africa	41	156	282
Netherlands Antilles	142	122	-14
Trinidad and Tobago	56	111	101
Guyana	74	104	41
Costa Rica	3	103	3,320
Poland	NA	94	NA
Egypt	29	41	41
United Arab Emirates	30	38	27
Guatemala	11	3	-72
Vietnam	NA	3	NA

Table 3: CDIA Ranked by Growth

	Ave(87-91)	Ave(99-03)	Growth (%)
Hungary	2	6,807	340,240
Peru	9	1,924	21,282
Sweden	11	1,268	11,225
Luxembourg	9	800	8,789
Cayman Islands	143	6,827	4,674
China	16	602	3,781
Argentina	128	4,924	3,747
Costa Rica	3	103	3,320
Colombia	25	794	3,050
South Korea	28	821	2,868
Thailand	33	918	2,724
Austria	26	645	2,381
Chile	230	5,704	2,378
Ecuador	9	212	2,253
Barbados	1,351	23,136	1,613
Mexico	218	3,293	1,413
Norway	32	423	1,229
British Virgin Islands	29	313	998
Panama	17	187	974
Ireland	1,173	11,763	903
Netherlands	1,188	10,316	768
Philippines	30	258	764
Japan	890	7,023	689
Germany	800	5,917	639
Italy	380	2,705	612
Malaysia	81	560	595
Hong Kong	522	3,134	500
Bermuda	1,769	9,823	455
Belgium	535	2,651	396
New Zealand	152	675	345
Brazil	1,521	6,557	331
Bahamas	1,872	7,738	313
Switzerland	978	4,039	313
France	1,456	5,634	287
South Africa	41	156	282
Indonesia	997	3,668	268
Portugal	126	460	265
United Kingdom	11,205	36,307	224
Denmark	35	110	219

United States	55,981	176,450	215
Papua New Guinea	86	269	212
Venezuela	81	247	205

Figure 1

Canadian Direct Investment Abroad by Sector: Industrial countries (99-03)

Energy & minerals 19%

Others 24%

Machinery 7%

Services 15%

Finance & insurance 35%

Figure 2

Canadian Direct Investment Abroad by Sector: Less Developed Countries (99-03)

Others 11%

Services 7%

Energy & minerals 27%

Machinery 2%

Finance & insurance 53%
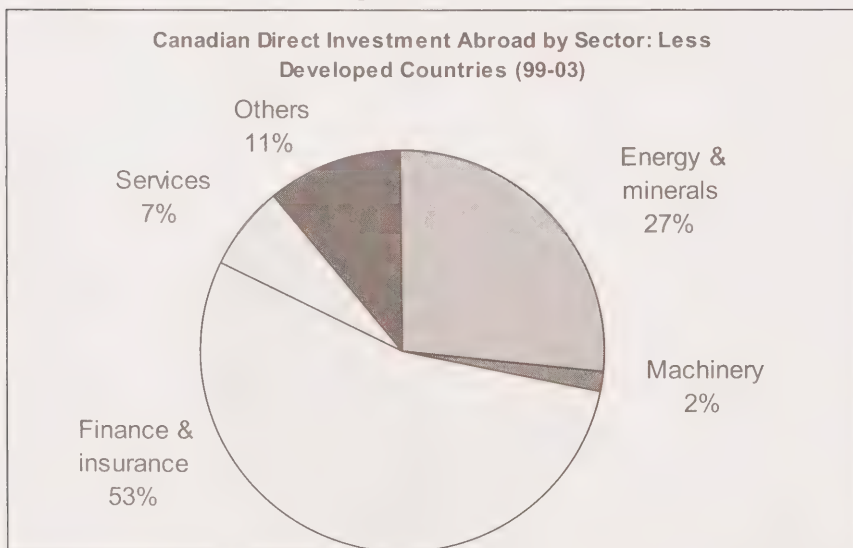
A Framework for Examining Determinants of FDI

The Lucas Paradox of the Neoclassical Model

Following Lucas (1990), assume a small open economy where output Y is produced using capital K and labour L.

1. $$Y_t = A_t F(K_t, L_t) = A_t K_t^{\alpha} L_t^{1-\alpha}$$

where A is the productivity parameter and the t subscript denotes time period t. If all countries share a common technology, perfect capital mobility implies the convergence of interest rates. The property of diminishing returns to capital implies that resources will flow from low-returns to high-returns countries. Hence, for countries i and j,

2. $$A_t f'(k_{it}) = r_t = A_t f'(k_{jt})$$

This is certainly not what we observe in the data. There is insufficient capital flowing from rich to poor countries to yield the convergence of interest rates across these countries. Lucas suggests that the lack of capital flows from rich to poor countries can be accounted for by looking at the existence of other factors such as the differences in human capital quality and the role of human capital externalities that might positively affect the productivity of capital. For instance, if the availability and quality of human capital positively affects the return to capital, FDI will flow to countries with large endowments of high quality human capital. In other words, Lucas essentially assumes that the productivity parameter is higher in rich countries (A_{rich} > A_{poor}) because of a human-capital externality. If human capital positively affects the rate of return to capital, investment will not flow to poor countries, as the rate of return there will be much lower than the neoclassical theory suggests. In equilibrium, workers can earn higher wages in rich countries than in poor countries, and there is no pressure on capital to flow to poor countries.

Generalizing this argument, the paucity of capital flows from rich to poor countries is explained by the existence of

other factors that affect the rate of return to capital, which in turn, influences the location decision of investment. Thus, the production function (1) can be rewritten as

3. $$Y_t = A_t F(K_t, Z_t, L_t) = A_t K_t^\alpha Z_t^\beta L_t^{1-\alpha-\beta}$$

where Z_t denotes the "missing factors" that account for the lack of capital. Thus, the true rate of return that accounts for the return differentials between countries i and j should be

4. $$A_t f'(k_{it}, z_{it}) = r_t = A_t f'(k_{jt}, z_{jt})$$

The following discusses the main components of z_{it} in turn. We modify the approach of Lucas (1990) to incorporate other determinants of FDI flows that are found in the literature. In particular we attempt to empirically substantiate three different theoretical motivations for FDI: market-seeking FDI, resource-seeking FDI, and efficiency-seeking FDI. In addition to these main motivations for FDI, we incorporate economic policy variables and measures of political risk into our analysis. We adopt the gravity model that has become the most widely employed empirical model for understanding the geographic distribution of economic relationships. The gravity model allows us to empirically examine the determinants of CDIA based on the above framework.

The economic determinants of FDI

The literature establishes three key motivations for FDI: market-seeking, resource/asset-seeking; and efficiency-seeking. The literature tends to find that macroeconomic stability, country risk and the policy environment are also important determinants of FDI. Table 4 lists these six determinants of FDI and provides some measures for each determinant. This sub-section discusses each determinant in turn.

   The key to understanding the determinants of FDI is to understand the interaction of FDI with its chief agent, the multinational enterprise (MNE). A MNE is a firm that controls and

manages production establishments located in at least two countries. These establishments may be horizontally or vertically integrated. Horizontally integrated establishments produce the same line of goods or services in each facility in different locations; vertically integrated establishments produce output in facilities that serve as an input for other facilities located across national boundaries. Consequently, FDI primarily emerges as a tool through which a MNE extends its territorial boundaries abroad. In other words, FDI is a mechanism for MNEs to maintain control over production activities outside their own national boundaries (Dunning and Rugman, 1985).

Table 4: Host country determinants of FDI

Determinants of CDIA	Variables measuring host country characteristics
1. Market-seeking	Market size: real GDP or population
	Wealth: per capita income
	Market growth: growth of real GDP
2. Resource/asset-seeking	Country resource endowment: the proportion of resource exports in total exports by country
3. Efficiency-seeking	Human capital: an education index is used to measure the quality of the workforce.*
4. Macroeconomic stability	Inflation
5. Country risk	Expropriation risk
	Rules of law regarding entry and operations
	Repudiation of contracts by government
6. Institution and policy environment	Infrastructure
	Taxation policy
	Investment agreements
	Trade commissioners

* NB: We do not have data on wages or efficiency wages so this proxy is used instead.

Market-seeking FDI

The importance of market-seeking as a determinant of FDI is self-evident. Various empirical studies have shown that the

size[2] and growth[3] of host country markets are among the most important determinants of FDI. The idea here is simple: large size and high economic growth increase the attractiveness of a country for a MNE, since a larger market size enables it to exploit economies of scale and earn higher profits.

Recently the market-seeking motive for FDI gained in importance with the opening of services industries to FDI. The bulk of FDI in services, which accounts for a rising share in overall FDI, is market-seeking almost by definition, as many services require face-to-face delivery.[4]

In the analysis that follows, we use real GDP to measure country size and real GDP per capita to measure wealth. We also construct a variable called "emerging", which combines country size with growth. Note that, as discussed below, the market size of the host country is also a determinant of FDI, as derived from the gravity model.

Resource/asset-seeking

Resource-seeking FDI is motivated by the availability of natural resources in host countries. This type of FDI was historically important and resource-seeking remains an important factor in attracting FDI for some developing and emerging market countries. The share of resource-seeking FDI in total world FDI flows, however, has decreased significantly. The primary sector's share in the outward FDI stocks of major investing countries was below 5 percent in the first half of the 1990s (UNCTAD, 1998).

---

[2] See Agarwal (1980), Brewer (1993), Lucas (1993), Bajo-Rubio and Rivero (1994), Wang and Swain (1995), Gastanga, Nugent and Pashamova et al (1998).

[3] See Agarwal (1980) and Lucas et al. (1993).

[4] Note there are four modes of services supply: cross-border trade, consumption abroad, commercial presence, and movement of natural persons. Commercial presence in services by definition is market-seeking FDI. Although complete data on the four modes of services are not available, available data on exports shows that Canadian services exports through commercial presence are twice as large as that through cross-border trade.

The declining share of resource-seeking FDI is partly due to the decline in relative prices of resources. Note, however, that resource-seeking remains an important motivation for Canadian firms investing abroad. The share of resource-seeking FDI flowing from Canada was 24 percent of total outward FDI in 2002—the same share as in 1989.[5] Given the continued importance of resource-seeking FDI to Canadian firms, we include a measure of resource potential in our analysis.

Efficiency-seeking

Efficiency-seeking (i.e., a desire to take advantage of differences in labour costs, technologies, and specialization) has emerged in recent years as an important motivation for FDI and has played an important role in the rise of geographically dispersed production chains. Although cost differences are important factors in other types of FDI, the competition for efficiency-seeking FDI flows is based entirely on cost differences between locations, the quality of infrastructure and business-related services, the ease of doing business and the availability of skills and technologies.

The naïve articulation of this motive is that firms invest in markets to exploit low wage workers. However, the empirical evidence is clear: labour productivity plays a key role in the efficiency-seeking motivation for FDI. Thus, it is low efficiency-wages that are important, not low wages *per se*. The "efficiency wage" concept combines the wage rate paid and the productivity of the workers. The evidence on this is compelling. Miller (1993), Wang and Swain (1995), Bajo and Rivero (1994), and Lucas (1993) provide evidence that labour costs themselves are not important determinants of FDI inflows. Therefore, we do not include wage levels as a determinant of CDIA, but we do include a measure of human capital, as a rough proxy for efficiency wages.

Macroeconomic stability

The institutional and policy environment is a key determinant of success in attracting FDI. This criterion includes economic and political variables that we discuss in turn. According to Brewer (1993), if market size is an important factor in explaining FDI, macroeconomic policies such as monetary and fiscal policies that stimulate or retard growth are also significant determinants of FDI.

Macroeconomic variables are known to indirectly impact FDI as they provide information about the expected state and stability of the economy. For example, in his survey article, De Mello Jr. (1997) refers to FDI inflows as being sensitive to balance of payments constraints and to general macroeconomic performance. Similarly Bajo and Rivero (1994) and Wang and Swain (1995) use macroeconomic variables such as inflation or real interest rates as proxies for the degree of macroeconomic stability and the expectations associated with an economy. Wang and Swain (1995) find evidence of significant negative effects of domestic interest rates on FDI inflow to China and Hungary over the periods 1978-1992 and 1968-1992 respectively. Empirical evidence provided by Bajo and Rivero (1994) shows a highly significant effect of inflation on FDI inflows.

The level and variability of the exchange rate also likely plays an important role in determining FDI inflows. However, the empirical evidence on exchange rates and FDI is decidedly mixed. Bajo and Rivero (1994) find an insignificant impact of the real effective exchange rate on FDI inflows. Empirical studies conducted by Wang and Swain (1995) and Blonigen (1997) have also generated mixed support for a link between exchange rates and FDI inflows. Blonigen (1997) acknowledges that empirical studies have had limited success establishing a link between exchange rate levels and FDI inflows. Lucas (1993) suggests the same, ascribing a residual role for exchange rate risk in determining the value of repatriated profits.

Although exchange rate movements impact firm dynamics and decisions, the relationship is complex and differs across firms. It is very difficult to find an empirical relationship be-

tween exchange rate movements and aggregate FDI. The inflation rate is the preferred variable that reflects macroeconomic stability and is found to be a significant determinant of FDI flows. In the empirical analysis reported below, we therefore include inflation as a criterion of macroeconomic stability.

Country risk

The absence of secure property and contractual rights discourages investment (Olson, 1982; Weingast, 1993). Knack and Keefer (1995) develop institutional indicators to measure the country risks facing investors. These institutional indicators include *expropriation risk* (measuring the risk of expropriation) *rules of law* (measuring whether there are established peaceful mechanisms for adjudication of disputes), *repudiation of contracts by government* (measuring contract enforcement), and *corruption in government and quality of bureaucracy* (measuring the general efficiency with which government services are provided). Knack and Keefer (1995) find that institutions that protect property rights are crucial to economic growth and investment. Countries that score low on these dimensions are likely to have less physical and perhaps human capital investment. The low capital stock of countries with poor institutions does not attract investment despite the possibility of high returns. On the other hand, high capital stock countries, benefiting from good institutions, continue to attract investment despite diminishing returns to investment.

A number of available measures provide information on the "Lack of Investor Protection" across countries.[6] There are two basic approaches to measuring "investor protection" or "country risk": *objective indicators* and *subjective indicators*. Objective indicators are based on observed institutions – for example, an objective measure of investor protection is the extent of formal protections of property rights a country has codified into the

---

[6] Export Development Canada, Transparency International and the World Economic Forum were also considered as possible data sources; however, the ICRG and KKM data were deemed to be most appropriate for this analysis.

legal system. Subjective indicators are based on survey data from firms and/or individuals either through polls or expert opinion.

Both types of indicators quantify risk characteristics and typically cover various aspects of investor protection and country risk. The indicators differ in their measurement approach, their focus and their coverage of years and countries. Both types of indicators are useful and have been widely used in the literature. However, the architects of the subjective indicator point out:

> *"...we rely exclusively on subjective or perceptions-based measures of governance when constructing our aggregate governance indicators. The primary reason for this choice is that for many of the key dimensions of governance, such as corruption or the confidence that property rights are protected, relevant objective data are almost by definition impossible to obtain, and so there are few alternatives to the subjective data on which we rely."* (Kaufmann, Kraay and Zoido-Lobatón (2004, p. 19)

The interested reader may refer to the Appendix for more information on the differences between subjective and objective indicators. Both types of indicators are used in this analysis.

Daniel Kaufmann, Aart Kraay, and Massimo Mastruzzi have constructed indicators covering 199 countries but only for 1996, 1998, 2000, and 2002. The data are further described in the World Bank Policy Research Working Paper 3106 titled "Governance Matters III: Governance Indicators for 1996-2002". We use and report results for the indicator for Rule of Law; the measures for Regulatory Quality and Control of Corruption are used in the analysis described below although the results using these variables are not reported.

The second group of country risk indicators is produced by the PRS Group; these indicators cover 145 countries over the period 1984-2003. The indicators for Investment Profile, Law and Order, and Corruption are used.

Both groups of indicators come in point estimates.

Institutions and the policy environment

Infrastructure: The existing infrastructure in a country affects the cost of doing business and therefore potentially plays a role in firms' decisions to locate in a given jurisdiction. Data on infrastructure is difficult to obtain for a broad cross section of countries but we attempt to include some measures such as telephone infrastructure.

Taxation/regulation policy: Taxation and regulatory policy potentially have important implications for firm investment decisions. For example, UNCTAD (2004) lists privatization as a leading determinant of recent FDI flows into Latin America. This is an extreme example of how regulation affects investment location decisions. There is also a very well-developed literature looking at the impact of taxation rates on firm investment decisions. Beaulieu, McKenzie and Wen (2005) provide a brief survey of this literature. They point out that effective tax rates are the appropriate measure of tax structure of a given jurisdiction. However, comprehensive cross-country data on effective tax rates do not exist. In the analysis below, we include data on the marginal corporate tax rate by country over time.

Trade policy: Trade policies can affect the incentives for FDI in many ways. A sufficiently high tariff may induce tariff-jumping FDI to serve the local market. Other types of import barriers can have the same effect. It is no coincidence that Japanese automobile manufacturers began producing in the United States following the imposition of so-called "voluntary export restraint" agreements limiting the number of automobiles that could be shipped from Japan. Such quid pro quo investments are motivated by the belief that the added cost of producing in the foreign market is more than compensated by the reduced probability of being subjected to new import barriers on existing exports to that market. However, the gains to the recipient country from this type of FDI may be limited. FDI attracted to protected markets tends to take the form of stand-alone production units, geared to the domestic market and not competitive for

export production. High tariffs on imported raw materials and intermediate inputs can further reduce international competitiveness, especially if local inputs are costly or of poor quality. Furthermore, if trade is subsequently liberalized, tariff-jumping investment would likely be among the first candidates to be rationalized, as it is not competitive for export production. In this case, trade policy and investment policy offset each other thereby reducing the potential benefits of trade and investment liberalization.

On the other hand, trade and investment policy can complement and reinforce each other, leading to greater trade and investment flows. For instance, the increases in efficiency-seeking investment due to investment liberalization and low tariffs could stimulate trade in intermediate inputs, which has proven to be one of the driving forces behind the spectacular growth of merchandise trade we have witnessed in the past several decades. In addition, if FDI stimulates economic growth in the host country, as appears to be the case, the result will be an increase in demand for imports, including from the source country.

In the analysis that follows we include a measure of the trading relationship between countries and consider the impact of trade on CDIA.

Investment Agreements: Canada has signed 43 investment agreements, in part to provide some level of investment protection in higher risk jurisdictions. We examine whether investment agreements affect CDIA.

The recent literature examining the impact of bilateral investment treaties (BITs) on international investment flows provides somewhat mixed results. Neumayer and Spess (2004) examine the impact of BITs on FDI flows into developing countries; they find that the number of BITs increases FDI flows to developing countries. They argue that BITs function as substitutes for good domestic institutional quality. That is, countries with poor domestic institutional quality gain the most from signing bilateral investment treaties. On the other hand, Mary Hallward-Driemeier (2003) examines twenty years of FDI flows

from the OECD to developing countries and she finds little evidence that BITs have stimulated additional investment.

Tobin and Rose-Ackerman (2004) empirically examine whether BITs affect investment flows from 176 countries and include various indicators of government performance, investment rates, social indicators and investment treaties. The authors adopt a panel approach using data from 1959 (the year of the first BIT) through 2000. The authors examine the evidence in two ways: they look at FDI flows in general; and they look at bilateral US data.

For the general case, they look at the relationship between a country's share of total world FDI and the number of BITs the country has ratified. The BIT variable is measured as the cumulative number of BITs that a country has signed using five year averaging to avoid year-to-year variation in the number of BITs from 1975 to 2000. The authors find that the number of BITs a country has signed has no statistically significant effect on FDI flows.

When they look at US bilateral data, they find a negative relationship between US FDI and BITs. This result may reflect the fact that countries with low levels of FDI from the US are clamouring to sign agreements in an attempt to attract more US FDI. In a related vein, Blonigen and Davies (2004) find that bilateral tax treaties between the US and its partners do not increase US FDI to those countries.

Empirical Analysis

The gravity model

The most widely employed and successful model for understanding the geographic distribution of economic relationships is the gravity model. We develop a version of this model to empirically examine the economic and commercial determinants of Canadian FDI, using the variables identified in the literature review. We discuss two estimating approaches: 1) a cross section estimation; and 2) a pooled cross-section and time series estimation.

Cross section analysis

In principle, it is possible to estimate a cross-section model to examine Canada's FDI abroad. The advantage of this approach is that the data requirements are reduced. The downside of using a cross-section approach is that we exclude variables that change over time.

Since we are only considering Canada as the source country, we drop the "i" subscripts and estimate Canada's investment to country j. Thus, our estimating equation becomes:

5. $CDIA_j = \beta HostVariables_j + \varepsilon_j$

where $CDIA_j$ is Canadian (the source country) FDI to host country j. Our estimates of CDIA are obtained from Statistics Canada.

As mentioned, typically the gravity equation would include home, or source, country as well as host country characteristics, but in this case the home country characteristics are for one country, and hence there is no variation in the cross-country regression. Therefore, in a cross-section regression, the Canadian (source country) characteristics are not included; *HostVariables* includes only the characteristics of the host countries.

Cross-section and time-series panel approach

A more powerful estimating approach is to take advantage of the panel nature of the data. The panel data include variables that vary over time as well across countries. Again, since we are only considering Canada as the source country, we continue to suppress the "i" subscripts and estimate Canada's investment to country j at time t. Thus, our estimating equation becomes:

6. $CDIA_{jt} = \beta HostVariables_{jt} + \alpha_j + \lambda_t + \varepsilon_{jt}$

where $CDIA_{jt}$ is Canadian (the source country) FDI to host country j at time t. We include the same dependent and explanatory variables as discussed above. In addition we include

a country fixed effect (α_i) – which controls for country characteristics that are not observed in the data; and we include a trend variable (λ_t).

We have constructed two datasets on cross-country CDIA and investment risk. We have a dataset that has data for four years: 1996, 1998, 2000 and 2002. We also have a dataset that covers 1987 to 2003.

For market-seeking FDI, we use real GDP (constant 1995 USD) as a proxy for market size. These data are from the World Development Indicators constructed by the World Bank Group. The data cover 208 countries over the period 1960-2002, although not all years have data for all countries.

Inflation is used as a measure of a country's macroeconomic stability. These data come from the International Monetary Fund's September 2004 World Economic Outlook Database, and cover 179 countries for 1980-2005.

For infrastructure we use the average cost of a phone call to the US from the World Development Indicators.

We include average corporate taxation data, which are from the World Tax Database compiled by the University of Michigan.

Before looking at the regression results, it is worth looking at some scatter plots of CDIA and some of the key variables. Figure 4 presents a scatter plot of CDIA in 2002 against the distance to recipient countries. We can see that there are a large number of countries grouped together with less than $10 billion dollars in CDIA. However, we still see that the regression line has a negative slope: the destination of CDIA is negatively associated with distance. Recall that the United States accounts for over 47 percent of CDIA and that the United Kingdom accounts for 9.7 percent. Figure 5 re-draws the relationship between CDIA and distance after excluding these two countries and again we see a negative relationship between CDIA and distance. This is consistent with what we expect from the literature.

Figure 6 looks at the simple linear bivariate relationship between CDIA and corporate tax rates. Again, excluding the United States and the United Kingdom, CDIA is negatively related to corporate tax rates. This is what we expect to see based

on our discussion of the trends and patterns of CDIA. CDIA
flows are generally larger to low tax jurisdictions. Notice that
Barbados (BRB) receives the most CDIA among these countries
but shows a very high corporate tax rate. It is important to note
that this is an aggregate tax rate and that the tax rate on the fi-
nancial sector (where Canadian CDIA is concentrated in that
country) is 5 percent. If we exclude Barbados from the diagram,
or code the corporate tax rate at 5 percent in that country the
slope of the trend line would be even steeper in absolute terms.

Finally Figure 7 presents the relationship between CDIA
and "rule of law". As we can see from this diagram the trend
relationship is positive. That is, the distribution of CDIA across
countries is positively correlated with the "rule of law" in those
countries. As expected, CDIA is higher in countries with
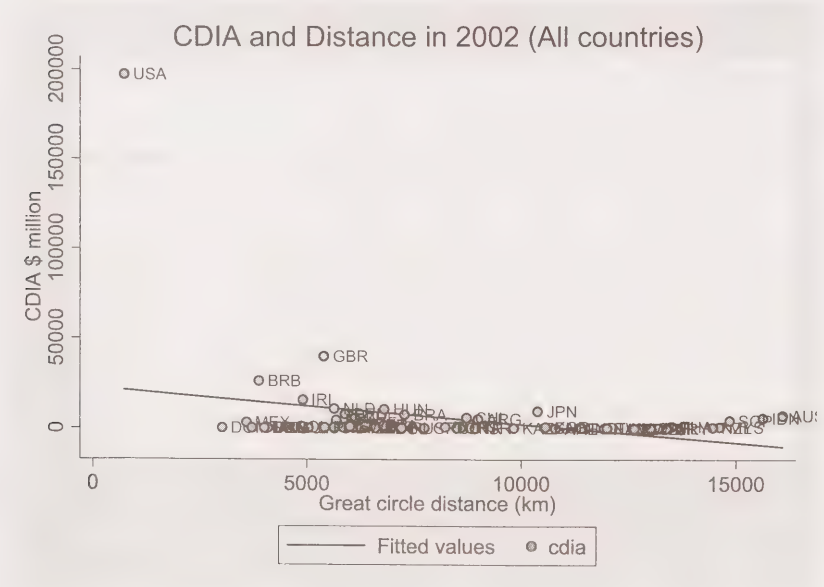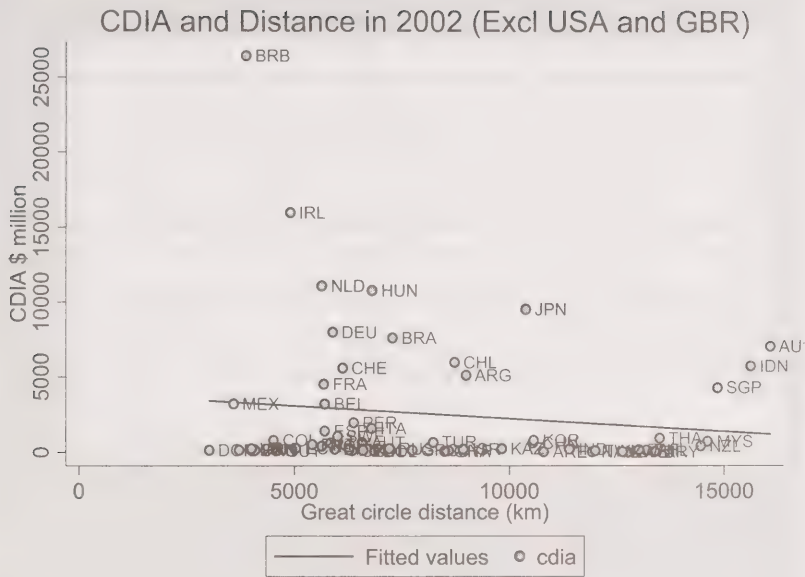stronger legal institutions.

Figure 4

CDIA and Distance in 2002 (All countries)

Figure 5

CDIA and Distance in 2002 (Excl USA and GBR)

Figure 6

CDIA and Corporate Tax Rate in 2002 (Excl. USA and GBR)

131
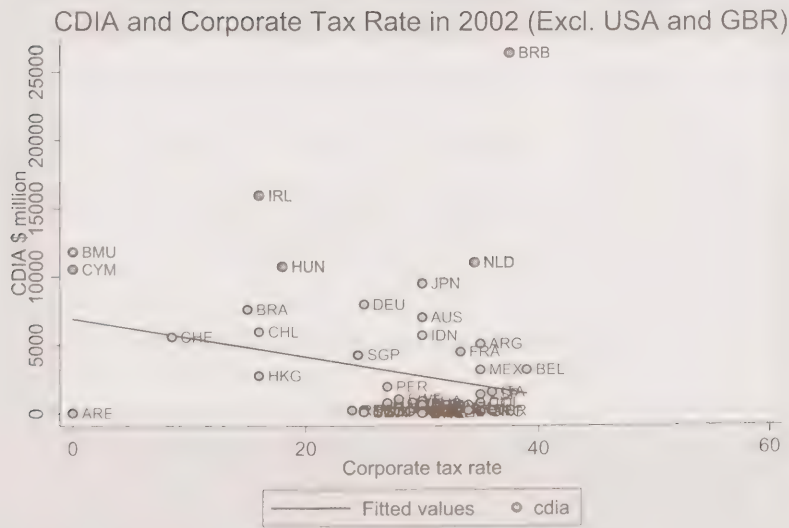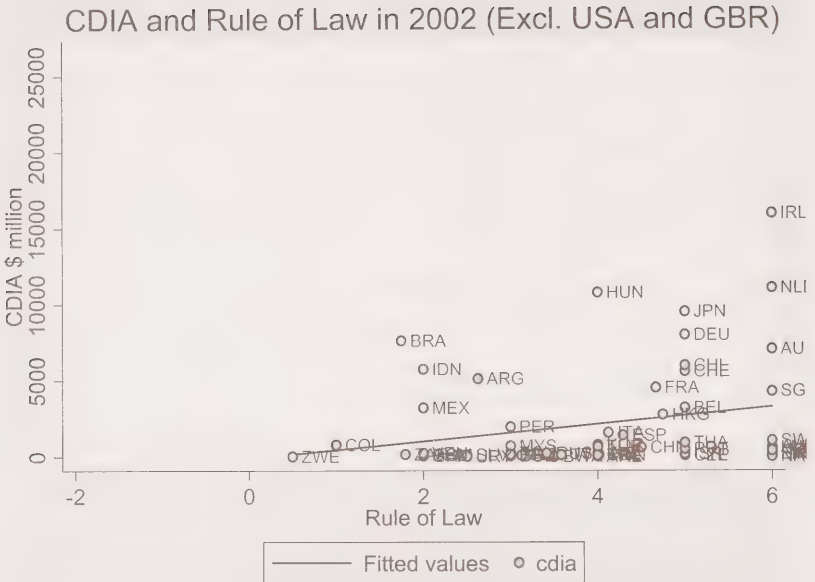
Figure 7

CDIA and Rule of Law in 2002 (Excl. USA and GBR)

The patterns of CDIA in these scatter plots reveal some anticipated correlations. However, it is important to see whether these correlations hold in a multivariate framework. As shown below, the results from multivariate regressions are similar.

Cross Section Results

Table 5 reports four columns of results from estimating Equation 7 based on 2000 data only, using different versions or specifications of the gravity model. In addition to a common set of variables, column 1 includes per capita income, column two includes rule of law, column 3 includes the education index and total trade, and column 4 includes rule of law and total trade.

The effect of GDP on CDIA is positive and statistically significant in all four specifications, while the estimated effect of distance on CDIA is negative and statistically significant in all four estimations. These results are robust and consistent both with our expectations and with the results found in the literature.

The coefficient on the corporate tax rate is negative as ex-

pected and is statistically significant in all four specifications; thus, a low corporate tax rate in the host countries provides a powerful and effective incentive to Canadian investment.

The abundance of resources in the host countries is positively related to CDIA in all four specifications. This reflects Canada's major presence in the global resource market. Resource sector investments account for roughly one quarter of total CDIA.

The table also reports the estimation results of the impact that signing a Foreign Investment Protection Agreement (FIPA) has on CDIA. The variable "FIPA" is equal to 1 if a country has a FIPA with Canada and zero otherwise. In all four specifications, a FIPA does not have a statistically significant effect on CDIA.

Table 5: Cross section regression results

Dependent Variable: LOG(CDIA)

Variable	(1)	(2)	(3)	(4)
C	-5.4957	-2.1189	-2.87686	-0.22796
Log(GDP)	0.7243*	0.8282*	0.5702*	0.5079*
Log(Distance)	-0.8666**	-1.1495*	-0.9323**	-0.9978*
Log(GDP Per Capita)	0.4205*			
Education Index			2.2418	
Law		0.5012*		0.4762**
Corporate Tax	-0.1025*	-0.1055*	-0.1138*	-0.1032*
Resources	0.0547*	0.0551*	0.0498*	0.0594*
FIPA	0.1401	0.168558	-0.1254	0.225543
Trade Commissioners	0.0091	-0.00069	-0.0008	-0.00637
Log(Total Trade)			0.3026**	0.3572*
R-squared	0.5879	0.5812	0.5816	0.5974

Note: * significant at the 5 percent level; ** significant at the 10 percent level.

The same is true for another policy instrument, "trade commissioner services", measured as the number of Canadian trade commissioners assigned to a country plus the number of locally engaged staff. While trade commissioners have no mandate to promote outward investment, this variable is included in order to test the *indirect* effect of trade promotion activities on outward investment. No such effect is, however, observed: in all four specifications, the estimated coefficients for "trade commissioner services" show no statistical significance.

As mentioned, the first specification also includes per capita income. As shown, our results find the estimated coefficient for per capita income to be greater than zero, implying that CDIA is positively correlated with GDP per capita.

The second specification includes the rule of law variable while the fourth specification incorporates total trade along with rule of law. Rule of law is included in order to test the importance of institutions in the host country in attracting CDIA. The results show a strong positive effect of rule of law on CDIA. Canadian investors are more likely to invest in countries with good domestic institutions, particularly sound legal systems.

The third column of Table 5 includes an education index to test the role of human capital in determining the pattern of CDIA. The estimated coefficient is positive but statistically insignificant suggesting that, *ceteris paribus*, the level of human capital in the host country is not a critical factor in attracting CDIA.

Lastly, the third and fourth columns explore the correlation between CDIA and bilateral trade with host countries. The estimated coefficients are positive and statistically significant. This suggests that Canadian trade and investment are complements, not substitutes as some economists suggest.

Pooled time series and cross section results

Table 6 presents results from estimating a pooled regression based on the gravity model presented in Equation 8. The first two columns present the results from estimating the model over all years (1987 to 2002). Columns 3 to 5 report the results from estimating the model over a sub-sample of years: 1987-91 in column 3; 1992-96 in column 4; and 1997-02 in column 5. All five regressions include distance, rule of law, the variable "emerging" (which measures the effect of economic growth and size of GDP), real GDP per capita, inflation, the corporate tax rate and a time trend. The simulation reported in column 1 also includes a measure of infrastructure based on the availability of telephones. The telephone variable restricts the number of observations in the model. All variables in the regressions are log values.

Table 6: Pooled regression results

	(1)	(2)	(3)	(4)	(5)
	1987-02	1987-02	1987-91	1992-96	1997-02
Distance	-0.892	-0.88	-0.856	-0.816	-0.95
	(6.15)**	(7.04)**	(3.60)**	(3.57)**	(4.98)**
Rule of Law	0.387	0.366	0.547	0.484	0.134
	(1.70)	(2.01)*	(1.41)	(1.19)	(0.54)
Emerging (growth * size)	0.594	0.539	0.507	0.472	0.654
	(12.59)**	(13.60)**	(5.73)**	(6.78)**	(11.12)**
Real GDP per capita	-0.067	0.134	0.164	0.068	0.168
	(0.73)	(1.99)*	(0.98)	(0.59)	(1.76)
Inflation	-0.245	-0.119	0.002	-0.187	-0.097
	(3.73)**	(2.52)*	(0.03)	(2.15)*	(1.08)
Telephone	1.98				
	(1.06)				
Corporate tax	-0.052	-0.059	-0.061	-0.055	-0.064
	(3.92)**	(5.96)**	(4.01)**	(3.37)**	(3.59)**
Time trend	0.121	0.101	-0.044	0.159	-0.003
	(5.90)**	(6.53)**	(0.40)	(2.27)*	(0.05)
Constant	10.757	9.667	9.506	9.337	11.243
	(6.23)**	(7.06)**	(3.55)**	(3.72)**	(5.07)**
Observations	444	567	151	221	195
R-squared	0.45	0.43	0.37	0.34	0.54

The t statistics are given in the parentheses.
* significant at 5 percent level; ** significant at 1 percent level.

The estimated effect of distance on CDIA is negative and statistically significant in all five regressions. This extremely robust result is consistent with our prior expectations and with the previous literature. The "rule of law" variable is positive in all columns but not statistically significant except in column 2 where all years are included and telephone is excluded. The constructed variable "Emerging" is positive and statistically significant in all five versions of the model. Again, this is a robust result with the

expected sign. Real GDP per capita is not statistically significant. Inflation is negative as expected and is statistically significant in some cases. Corporate tax rate is negative as expected and is statistically significant. This is a robust result.

Table 7 presents results using pooled cross section and time series data to examine the impact of signing a FIPA on CDIA.

Table 7: Pooled regression results including FIPA

	(1)	(2)	(3)
Distance	-0.892	-0.904	-0.883
	(6.15)**	(6.22)**	(7.07)**
Rule of Law	0.387	0.453	0.417
	(1.70)	(1.93)	(2.25)*
Emerging (growth * size)	0.594	0.574	0.518
	(12.59)**	(11.49)**	(12.45)**
Real GDP per capita	-0.067	-0.093	0.112
	(0.73)	(0.97)	(1.61)
Inflation	-0.245	-0.234	-0.113
	(3.73)**	(3.56)**	(2.36)*
Telephone	1.98	1.972	
	(1.06)	(1.07)	
Corporate tax	-0.052	-0.051	-0.057
	(3.92)**	(3.87)**	(5.74)**
Time trend	0.121	0.128	0.109
	(5.90)**	(6.06)**	(6.89)**
FIPA		-0.41	-0.484
		(-1.76)	(2.25)*
Constant	10.757	11.039	9.855
	(6.23)**	(6.28)**	(7.13)**
Observations	444	444	567
R-squared	0.45	0.45	0.44

The t statistics are given in parentheses.
* significant at 5 percent level; ** significant at 1 percent level.

The first column in Table 7 recreates column 1 of Table 6 for comparative purposes. Column 2 adds the FIPA variable, and column 3 omits the telephone infrastructure variable, as this increases the number of observations. As can be seen in columns 2 and 3, the FIPA variable has a negative sign, contrary to expectations.

Discussion and conclusions

This paper provides an empirical analysis of trends and patterns in Canadian Foreign Direct Investment Abroad (CDIA). The paper provides an overview of CDIA from 1987 to 2003, presents a survey of the literature on determinants of FDI, and integrates variables identified in this literature as being significant determinants of FDI (such as economic growth and size and measures of investor protection) into a gravity model framework. An extensive dataset of host country attributes was compiled by the authors and is used in the empirical work.

The results show that CDIA is mainly driven by economic fundamentals such as the size of the host country's market, natural resources, and distance from Canada. The results also suggest that domestic policy features in host countries, such as a good legal infrastructure and low corporate taxes are important determinants of the pattern of CDIA.

The effects of FIPAs in facilitating CDIA are inconclusive or indeterminate at best. There is no significant, positive relationship between the presence of a FIPA and the distribution of CDIA; in a number of regressions, the sign is actually negative. This result is consistent with the recent literature examining the impact of bilateral investment agreements on FDI. However, this result may be due to countries beginning to improve their domestic institutions before the FIPA itself is signed, so that the improvement and signing do not occur in the same period. A second possible explanation is that Canada has not signed FIPAs with the "right" countries. Most FIPAs that Canada has signed are with small far-off countries with poor investment opportunities, which may explain why FIPAs have not been successful at promoting CDIA to partner countries. A third possible

explanation is that the economic model used to determine the pattern of CDIA does not fully describe investors' behaviour.

It is hoped that the research presented here will contribute to a discussion on policy direction. The results offer a perspective on the following two policy statements:

> The Canadian economy also benefits from outward investment, which contributes to competitiveness and to increased R&D, and leads to technology transfers and spillovers to our economy. We must support such international business appropriately with the right rules-based framework, both within the receiving country and between Canada and that country: the confidence to invest and expand must be based on assurances of predictable and equitable treatment.
> (Canada's International Policy Statement, 2005, p.10).

> Emerging markets frequently offer production efficiencies, but are sometimes coupled with challenges such as limited intellectual property (IP) protection and restricted market access. We must decide if Canada's current array of policy instruments and business development tools is sufficient to meet the existing opportunities, challenges and competition we face in the marketplace. Is enough being done to encourage outward as well as inward investment?
> ("Developing ITCan's Emerging Markets Strategy")
> http://www.itcan-cican.gc.ca/em_back-en.asp (last updated 2004-11-18).

The first statement points out that the federal government must support international business with a rules-based framework. There is fairly strong evidence, both here and in the literature, that a host country's legal framework is an important determinant of FDI flowing into that country. However, it is not clear from the evidence presented here, nor from the literature, that bilateral investment treaties such as FIPAs will provide a substitute for sound domestic legal institutions.

Bibliography

Agarwal, Jamuna P. "Determinants of Foreign Direct Investment: A Survey," *Weltwirtschaftliches Archiv* 116, (1980), pp. 739-773

Bajo-Rubio, Oscar, and Simón Sosvilla-Rivero. "An Econometric Analysis of Foreign Direct Investment in Spain, 1964-89," *Southern Economic Journal*, 61, 1, (July 1994), 104-120.

Baldwin, R. E. and A. J. Venables (1995) "Regional Economic Integration," in *Handbook of International Economics*, eds. G. Grossman and K. Rogoff.

Bauman, Renato. "Integration and Trade Diversion," *Cepal Review*, 51, (December, 1993), pp. 133-147.

Beaulieu, Eugene, Kenneth J. McKenzie and Jean Francoise Wen (2004) "Do Taxes Matter for Firm Location? Evidence from Canadian Provinces," The University of Calgary Department of Economics Working Paper 2005-07.

Bergara, Mario, Witold J. .Henisz, and Pablo T. Spiller. "Political Institutions and Electric Utility Investment," *California Management Review*, 40, 2, (Winter, 1998), pp. 18-35

Bhagwati, Jagdish. "Agressive Unilateralism: An Overview," in King, Philip ed. *International Economics and International Economic Policy - A Reader*, San Francisco, California: McGraw-Hill, Inc., 1995.

Blonigen, Bruce A. and Ronald B. Davies "Do Bilateral Tax Treaties Promote Foreign Direct Investment?" mimeo University of Oregon.

Blomström, Magnus, and Ari Kokko. "Regional Integration and Foreign Direct Investment," *National Bureau of Economic Research*, Working Paper 6019, April 1997

Bohn, Henning, and Robert T. Deacon. "Ownership Risk, Investment, and the Use of Natural Resources," unpublished manuscript, Department of Economics, University of California at Santa Barbara, California, July 1997.

Brewer, Thomas L. "Government Policies, Market Imperfections, and Foreign Direct Investment," *Journal of International Business Studies*, (First Quarter 1993), pp. 102-120.

Burton, F. N., and Hisashi Inoue. "A Country Risk Appraisal Model of Foreign Asset Expropriation in Developing Countries," *Applied Economics*, 19, (1987), pp. 1009-1048

Carmichael, Benoit, J. C. Cosset, and K. P. Fischer. "Political Risk in Latin America Stock Markets: a Rational Expectations Approach," *Working Paper 92-54*, Université Laval, Québec, August 1992

Chase, Carmen D., James L. Kuhle, and Carl H. Walther. "The Relevance of Political Risk in Direct Foreign Investment," *MIR*, 28, 3 (1988), pp. 31-38

Cuthberson, K., S. Hall, and M. D. Taylor. *Applied Econometric Techniques*, New York: Phillip Allen, 1992

Diamonte, Robin L., John M. Liew, and Ross L. Stevens. "Political Risk in Emerging and Developed Markets," *Financial Analysts Journal,* (May/June 1996), pp. 71-76.

Dunning, John H. *International Production and the Multinational Enterprise*. London: George Allen & Unwin, 1981

- - -, "Explaining the International Direct Investment Position of Countries Toward a Dynamic or Developmental Approach." *Weltwirtschaftliches Archiv* 117, (1981), pp. 30-64

- -, and Alan M. Rugman. "The Influence of Hymer's Dissertation on the Theory of Foreign Direct Investment," *American Economic Review, Papers and Proceedings*, 75, (1993), pp. 228-232.

Economist, The. "Sour Mercosur," *The Economist Newspaper Limited*, 352, 8132, (August 14th, 1999), p. 13.

Erb, Claude, Campbell R. Harvey, and Tadas E.Viskanta. "Political Risk, Economic Risk, and Financial Risk," *Financial Analysts Journal,* (November/December 1996), pp. 29-46

Ferris, Stephen, G.Rodney Thompson, and Cailin Valsan. "An Analysis of Foreign Direct Investment in Latin America: The Case of the Guyanese Economy," *Pakistan Economic and Political Review,* 35,1, (Summer 1997), pp.11-24.

Gastanaga, Victor M., Jeffrey B. Nugent, and Bistra Pashamova. "Host Country Reforms and FDI Inflows: How Much Difference do they Make?," World Development, 26, 7, (July 1998), pp. 1299-1314.

Hallward-Driemeier Mary. "Do Bilateral Investmetn Treaties Attact Foreign Direct Investment? Only a Bit… and They Could Bite," The World Bank Policy Research Working paper #3121.

Inter-American Development Bank. *Statistics and Quantitative Analysis Unit.* accessible at http://www.iadb.org/int/sta/english/staweb/stats.htm.

International Monetary Fund. *International Financial Statistics.* CDROM version, 1999.

Jaspersen, Frederick Z, Anthony H. Aylward, and Mariusz A. Sumlinski. "Trends in Private Investment in Developing Countries – Statistics for 1970-94," International Finance Corporation Discussion Paper Number 28, Washington: The World Bank, 1995.

King, Philip ed. *International Economics and International Economic Policy - A Reader,* San Francisco, California: McGraw-Hill, Inc., 1995.

Krugman, Paul R. "Is Free Trade Passe?" in King, Philip ed. *International Economics and International Economic Policy - A Reader,* San Francisco, California: McGraw-Hill, Inc., 1995.

Loree, David W., and Stephen E. Guisinger. "Policy and Non-Policy Determinants of U.S. Equity Foreign Direct Investment," Journal of International Business Studies, (Second Quarter 1995), pp. 281-299

Lucas, Robert E. B. "On the Determinants of Direct Foreign Investment: Evidence from East and Southeast Asia," World Development, 21, 3, (1993), pp. 391-406.

Mello Jr, Luiz R. "Foreign Direct Investment in Developing Countries and Growth: A Selective Survey," The Journal of Development Studies, 34,1, (October 1997), pp.1-34.

Neumayer, Eric and Laura Spess (2004) "Do bilateral investment treaties increase foreign direct investment to developing countries?" mimeo Department of Geography and Environment, London School of Economics and Political Science

Organization of American States. *Investment Agreements in theWestern Hemisphere: A Compendium*. Washington, D.C.: OAS Trade Unit, 1997

Ramcharran, Harry. "International Bank Lending to Developing Countries: An Empirical Analysis of the Impact of Country Risk." *Multinational Business Review*, (Spring 1999), pp.83-91.

Rivoli, Pietra, and Eugene Salorio. "Foreign Direct Investment and Investment Under Uncertainty," *Journal of International Business Studies*, (Second Quarter 1996) pp. 335-357.

Schneider, Friedrich, and Bruno Frey. "Economic and Political Determinants of Foreign Direct Investment," World Development, 13, 2, (1985), pp. 161-175.

Tobin, J. & Rose-Ackerman, S. (2004). *Foreign Direct Investment and the Business Environment in Developing Countries: the Impact of Bilateral Investment Treaties*. Yale Law School Center for Law, Economics and Public Policy Research Paper No. 293.

Torrisi, C. R. "The Determinants of Direct Foreign Investment in a Small LDC," *Journal of Economic Development*, 10, (1985), pp. 29-45.

Trefler, Daniel (1993) "Trade Liberalization and the Theory of Endogenous Protection: An Econometric Study of U.S. Import Policy," *Journal of Political Economy*. Vol. 101, No. 1. Pp. 138-160.

Tsai, Pan-Long. "Determinants of Foreign Direct Investment and its Impact on Economic Growth," Journal of Economic Development, 19, (June 1994), pp. 137-163.

United Nations Conference on Trade and Development. *World Investment Report 1998: Trends and Determinants*, Geneva, Switzerland: United Nations, 1998.

United Nations Economic Commission for Latin America and the Caribbean. *Policies to Improve Linkages with the Global Economy*. Santiago, Chile: United Nations, 1995

United Nations Economic Commission for Latin America and the Caribbean. *Foreign Investment in Latin America and the Caribbean*. Santiago, Chile: United Nations, 1998

Wang, Zhen Quan, and Nigel J. Swain. "The Determinants of Foreign Direct Investment in Transforming Economies: Empirical Evidence from Hungary and China," *Weltwirtschaftliches Archiv* 131, 2, (1995), pp. 359-381.

Yeats, Alexander J. "Does Mercosur's Trade Performance Justify Concerns About the Effects of Regional Trade Arrangements? YES," Policy, Planning and Research Working Paper, Washington: World Bank.

Appendix A: Measuring Investor Protection

Objective indicators

One widely used objective indicator is "International Country Risk Guide" (or ICRG) from the PRS Group.[7] The ICGR provides up-to-date country risk ratings for 140 countries in three broad categories: financial, political and economic risk. For our purposes, we are interested in the composite measure of political risk and/or some of the components of the political risk variable. These data are up-to-date and are available from http://www.countrydata.com/reports/. The **Political Risk Points by Component** is Table 3B: *This table provides the Political Risk Rating and all of its component points for all countries monitored by ICRG, allowing you to select the date of your choice. This rating assesses political stability by assigning risk points to a pre-set group of the political risk components: Government Stability, Socioeconomic Conditions, Investment Profile, Internal Conflict, External Conflict, Corruption, Military in Politics, Religion in Politics, Law and Order, Ethnic Tensions, Democratic Accountability, and Bureaucracy Quality. The lower the risk point total, the higher the risk, and the higher the risk point total the lower the risk.*

An historical and alternative version of the ICRG data also exists and is available from the same source. This database is known as IRIS and is available from 1982-97. The advantage of this database is twofold: it includes very specific measures of investment protection (repudiation of contracts by government, and risk of expropriation) and it goes back to 1982. The disadvantage is that the most recent data point available is 1997 – so it is not useful for doing current analysis. The IRIS Dataset was originally constructed in 1993 by Steve Knack and Philip Keefer for the IRIS Center at the University of Maryland, based on data obtained from the **International Country Risk Guide (ICRG)**. The dataset includes computed scores for six vari-

---

[7] Go to PRSGroup web page http://www.prsgroup.com/icrg/icrg.html. Or go to the link http://www.countrydata.com/datasets/

ables: corruption in government, rule of law, bureaucratic qual-
ity, ethnic tensions, repudiation of contracts by government, and
risk of expropriation. Knack produced subsequent issues of the
data for an ongoing series of working papers from the IRIS
Center. In its current form, IRIS-3 contains data for the period
1982-1997. In the present analysis, we use the ICRG data that
covers 1987 to 2003. The two investor protection variables (re-
pudiation of contracts by government, and risk of expropriation)
are included as components of the "investor profile" variable in
the ICRG data but are not reported separately.

There are scores of alternative country risk measures. The
ICRG is widely used in academic and policy analysis and is
relatively inexpensive. An alternative source is BERI: Business
Environment Risk Intelligence (http://www.beri.com/). This has
country risk data as well as Mineral Extraction Risk Assessment
and Quality of Workforce Index.

Subjective indicators

The **Worldwide Governance Research Indicators Dataset**
provides six subjective indicators on "governance" constructed
by researchers at the World Bank Institute (WBI).[8] The data-
base covers 199 countries and has six dimensions of governance
in four different years: 1996, 1998, 2000, and 2002. The authors
use econometric techniques to construct the six indicators from
several hundred individual variables measuring perceptions of
governance, drawn from 25 separate data sources constructed
by 18 different organizations. The authors assign individual
measures of governance to categories capturing key dimensions
of governance, and use an unobserved components model to
construct six aggregate governance indicators in each of the
four years. The six indicators are called: Voice and Account-
ability; Political Stability and Absence of Violence; Govern-
ment Effectiveness; Regulatory Quality; Rule of Law; and Con-
trol of Corruption.

---

[8] See Kaufmann, Kraay and Zoido-Lobatón (1999a,b and 2003).

Kaufmann, Kraay and Zoido-Lobatón (1999a,b and 2004) discuss various methodological issues, including the interpretation and use of the data given the estimated margins of error. They define the indicators as follows:

1. **Voice and Accountability:** *a number of indicators measuring various aspects of the political process, civil liberties and political rights. These indicators measure the extent to which citizens of a country are able to participate in the selection of governments. We also include in this category indicators measuring the independence of the media, which serves an important role in holding monitoring those in authority and holding them accountable for their actions.*

2. **Political Stability and Absence of Violence:** *several indicators which measure perceptions of the likelihood that the government in power will be destabilized or overthrown by possibly unconstitutional and/or violent means, including domestic violence and terrorism. This index captures the idea that the quality of governance in a country is compromised by the likelihood of wrenching changes in government, which not only has a direct effect on the continuity of policies, but also at a deeper level undermines the ability of all citizens to peacefully select and replace those in power.*

3. **Government Effectiveness:** *the quality of public service provision, the quality of the bureaucracy, the competence of civil servants, the independence of the civil service from political pressures, and the credibility of the government's commitment to policies. The main focus of this index is on "inputs" required for the government to be able to produce and implement good policies and deliver public goods.*

4. **Regulatory Quality:** *is more focused on the policies themselves. It includes measures of the incidence of market-unfriendly policies such as price controls or inadequate bank supervision, as well as perceptions of the burdens imposed by excessive regulation in areas such as foreign trade and business development.*

5.   **Rule of Law:** *several indicators which measure the extent to which agents have confidence in and abide by the rules of society. These include perceptions of the incidence of crime, the effectiveness and predictability of the judiciary, and the enforceability of contracts. Together, these indicators measure the success of a society in developing an environment in which fair and predictable rules form the basis for economic and social interactions, and importantly, the extent to which property rights are protected.*

6.   **Control of Corruption:** *measures perceptions of corruption, conventionally defined as the exercise of public power for private gain. Despite this straightforward focus, the particular aspect of corruption measured by the various sources differs somewhat, ranging from the frequency of "additional payments to get things done," to the effects of corruption on the business environment, to measuring" grand corruption" in the political arena or in the tendency of elite forms to engage in "state capture". The presence of corruption is often a manifestation of a lack of respect of both the corrupter (typically a private citizen or firm) and the corrupted (typically a public official or politician) for the rules which govern their interactions, and hence represents a failure of governance.*

There are several advantages of this dataset. First is the methodology used to construct an index from **subjective evaluations** of country governance. As mentioned above these subjective indices are drawn from 25 separate data sources. Not all of the data sources cover all of the indicators, or all of the years. However, the authors construct the indices for 199 countries using sophisticated econometric techniques. An additional advantage of these data is that they are easy to access and free.

Appendix B: Background on the gravity model

The gravity model gets its name from its intellectual genesis: the "Law of Universal Gravitation" proposed by Newton in the 17th century. Newton law modeled the attractive forces between two objects i and j as positively related to the mass of the two objects and negatively related to the distance between them. The gravity model was first applied to economic relations between countries by Jan Tinbergen in 1962.

The basic gravity model of economic interactions between countries is based on the Newton-Tinbergen framework:

$$5. \quad F_{ij} = G \frac{M_i^{\alpha} M_j^{\beta}}{D_{ij}^{\theta}}$$

where F_{ij} is the "flow" of trade or investment between country i and country j. Flows depend positively on a constant G; and on the mass—or economic size—of each country M_i and M_j; and depend negatively on the distance between the countries, D_{ij}. There are three parameters in this specification of the gravity model that measure the responsiveness of flows between i and j to country i's mass (or GDP) (α), country j's mass (or GDP) (β), and distance (θ). A nice feature of this model is that taking logs of the equation provides a linear expression that can be estimated using standard econometric techniques:

$$6. \quad \ln F_{ij} = C + \alpha \ln M_i + \beta \ln M_j - \theta \ln D_{ij} + \varepsilon_{ij}$$

In the present analysis, we build on this basic framework to estimate the salient determinants of Canadian outward FDI.

This basic gravity model has been applied to both international trade and investment and has been extended to include additional factors that affect flows of trade and investment between countries. For example, researchers have included common language, contiguity, remoteness, income per capita, and regional trade agreements. The gravity model does a very good job of explaining trade and investment patterns.

Foreign Affairs and International Trade Canada's Trade Model, Version 2.0

Evangelia Papadaki, Marcel Mérette, Yu Lan and Jorge Hernandez [*]

Introduction

In a rapidly changing international environment, shaped by the emergence of new economic powers and increasingly complex arrangements amongst companies doing business across borders, the stakes for a trading nation such as Canada with trade equivalent to over 70 percent of GDP are high. Policy makers, academics, and the business community are thus equally concerned to understand better the impact of new trade and integration agreements that may help shape the future prosperity of Canada.

The ability to make quantitative assessments of the potential implications of proposed international trade agreements in a bilateral, multilateral or unilateral context provides valuable insight and complements qualitative arguments. Since general equilibrium effects must be taken into account to fully understand the impact of major trade policy changes such as broad liberalization on a multilateral basis or through a deeper bilateral or regional free trade agreement, computable general equilibrium (CGE) models have been used extensively to analyze such policy initiatives. Such has been the case during the political and economic debates surrounding the various phases of the

[*] Evangelia Papadaki is a senior trade policy analyst with the Office of the Chief Economist (CSE), Foreign Affairs and International Trade Canada (ITCan). Yu Lan and Jorge Hernández are junior economists on contract by CSE. Marcel Mérette is an associate professor at the University of Ottawa. The views expressed in this report are entirely those of the authors and do not necessarily reflect those of Foreign Affairs and International Trade Canada.

149

European Integration, and closer to home in the cases of the Canada-USA and North-American Free Trade Agreements.

The Office of the Chief Economist in Foreign Affairs and International Trade Canada has thus developed computable general equilibrium modelling capacity in order to provide quantitative assessments of a variety of possible trade policies that Canada is considering or might wish to consider. This note describes the modifications introduced into the first version of the Department's trade model,[1] to allow it to capture dynamic effects of trade policies, in addition to the static effects captured by the initial version.

Specification of ITCan CGE Model, Version 2.0

The ITCan CGE model Version 1.0 introduced the innovative feature of disaggregating Canada into six regions, therefore allowing for a regional as well as economy-wide assessment of hypothetical trade policies. Indeed Canada's recent experience has demonstrated that free trade agreements can have different effects on national and provincial trade. As the United States is Canada's most important international commercial partner, the first priority has been to develop a model that can assess possible policies that directly affect the Canada-US trading relationship. Version 1.0 thus fully models the United States. The Rest of the World is also fully modeled as one aggregate region.

Version 1.0 has been used to analyse various scenarios of closer integration between Canada and the United States, ranging from the adoption of a common external tariff, to the more ambitious full economic union, and demonstrated the differential impact of these trade policies on the regions of Canada.

Version 1.0 is, however, subject to certain limitations. In particular, it is a static model with fixed capital. As such, it compares two equilibria: the baseline equilibrium prior to pol-

---

[1] ITCan CGE model, Version 1.0, is described in Evangelia Papadaki, Marcel Mérette, Yu Lan and Jorge Hernández, "The International Trade Canada Trade Model, Version 1.0", *Trade Policy Research 2004*, (Department of Foreign Affairs and International Trade: Ottawa, 2004): 247-264.

icy reform and the equilibrium achieved after the implementation of the hypothetical policy change. The adjustment from one market-clearing equilibrium to another is instantaneous, as static models do not explicitly model time; the results are interpreted as long-term effects. No account is taken of possible incentives for changes in capital accumulation and investment in the post-reform environment, nor is account taken of the adjustment and transitional process towards the new long run equilibrium. However, since the number of years between the baseline and the long-run equilibrium is thought to be large, understanding the adjustment and transitional dynamics is important for policy purposes.

ITCan CGE model Version 2.0 has been developed to shed light on the latter questions. Version 2.0 is a dynamic model. Such models explicitly introduce time and consequently allow us to model the adjustment process. They also allow for capital accumulation and investment responses to trade policy reform. Generally speaking, allowing for capital accumulation tends to strengthen the response to trade liberalization as compared with static models.

There are two general ways to introduce dynamics into CGE models: a) the recursive approach, and b) the inter-temporal maximization approach.

Recursive CGE models solve for a sequence of static equilibria, connected with each other through capital accumulation. Investment in these models is either exogenous and equal to savings in any given time period or endogenous but dependent on actual rather than future rates of return to capital. Recursive dynamic models are popular because they present a rather simple way of introducing dynamics into otherwise static CGE models. Examples of recursive dynamic models include the GREEN model, (Burniaux et al., 1992), the DART model (reference), and the GTAP-Dynam (reference) model.

Dynamic models that use the inter-temporal maximization approach assume that economic agents exhibit rational, optimizing behaviour, and take into consideration the state of variables both at present and in the future in their decision-making. Savings and investment are not only endogenous but they are de-

rived by a decision-making process that is inter-temporal in nature. Inter-temporal dynamic models are significantly more complex because, instead of a sequential solving method, they involve a dynamic system that is interdependent over time and has to be repeatedly solved forward or solved simultaneously. Due to the complexity, these "fully" dynamic models have to compromise in terms of regional or sectoral detail to remain tractable. Thus, a number of inter-temporal dynamic models are one-country models (Pereira and Shoven, 1988; Bhattacharaya, 1996). The RICE model (Nordhaus and Young, 1996), though a multi-country fully dynamic model, does not incorporate sectoral disaggregation.

The ITCan CGE model Version 2.0 uses the latter inter-temporal maximization approach. The methodology developed is inspired by the seminal paper of Mercenier and Michel (1994) and later expanded by Lavoie et al. (2001). Version 2.0 maintains the unique feature of Canadian regional desegregation, but due to the computational complexity of the dynamic model reduces the number of industrial sectors to nine from the twenty four sectors of Version 1.0.

The Theoretical Framework of Version 2.0

In the model, we first define different commodity sets. Sectors of activity are identified by s and d, with S representing the set of all industries so that s, d =1,..,S. Regions are identified by indices i and j, with W representing the set of all regions so that i, j=1,..,W. In an inter-temporal, multi-region, multi-sector framework, it is necessary to keep track of trade flows over time by their geographical and sectoral origin and destination. Thus, a subscript $isjdt$ indicates a flow originated in period t and sector s of region i with industry d of region j as recipient. Since it will be necessary more than once to aggregate variables with respect to a particular subscript, to avoid unnecessary proliferation of symbols, occasionally we substitute a dot for the subscript on which aggregation has been performed; for instance, $c_{.siv}$ is an aggregate of c_{jsit} with respect to the first subscript.

Households

In each region, there is a representative household living infinitely and maximizing its utility. The household in each region chooses consumption and investment levels that maximize its utility. The household can borrow or lend from international financial markets. Through an inter-temporal decision-making process, the representative household in each region choose its aggregate levels of consumption and investment path over time and for each period t. Once optimal consumption at each time is determined, it is allocated among different sectors and subsequently over different geographic origins.

These distinct allocation steps are described for the steady state in greater detail below.

(a) The Inter-temporal Optimization Problem

The inter-temporal decision problem of the representative household of region i is to maximize:

$$\int_0^\infty e^{-\psi t} \frac{C_{it}^{1-\gamma}}{1-\gamma} dt , \tag{1}$$

subject to:

$$\dot{F}_{it} = \psi F_{it} + \left[w_{it} L_{it} + r_{it} K_{it} - PC_{it} C_{it} - PI_{it} I_{it} \right] \quad \dot{F}_{i\infty} = 0 \tag{2}$$

$$\dot{K}_{it} = I_{it} - \delta_i K_{it} ; \qquad \dot{K}_{it} = 0 \tag{3}$$

where:

ψ : Rate of time preference,

C_{it} : Total consumption of the household living in country i at time t,

γ : Inverse of the inter-temporal elasticity substitution,

F_{it} : Stock of foreign assets held by country i at time t,

$\dot{F}_{it}$: Change in the stock of foreign assets held by country i at time t,

w_{it} : Nominal wage rate in country i at time t,

L_{it} : Labour supply in country i at time t,

r_{it} : Nominal rental rate on capital in country i at time t,

K_{it} : Capital stock in country i at the beginning of time t,

PC_{it}: Consumption price index in country i at time t,

PI_{it} : Investment price index in country i at time t,

I_{it} : Total investment by the household of country i at time t,

δ_i : Rate of capital depreciation in country i ;

We borrow from a methodology developed by Mercenier and Michel (1994), and later expanded by Lavoie et al., (2001), that permits the transformation of continuous, infinite-horizon time into discrete, finite-horizon time, which can then be re-solved numerically. In the discrete, finite-horizon formulation time horizon can be aggregated into v time intervals or periods[2].

154

The first order conditions of the discrete finite horizon problem[3] are:

$$\log\left(C_{iv-1}\middle/ C_{iv}\right) = \frac{1}{\gamma}\log(PC_{iv}/PC_{iv-1}) \tag{4}$$

$$I_{iV} = \delta_i K_{iV} \tag{5}$$

$$F_{iV} = \frac{1}{\psi}\left[PC_{iV}C_{iV} + PI_{iV}I_{iV} - w_{iv}L_{iV} - r_{iv}K_{iV}\right] \tag{6}$$

$$PI_{iv\,1} = \frac{1}{1+\psi\Delta_v}\left[\Delta_v r_{iv} + (1-\delta_i\Delta_{iv})PI_{iv}\right];\ \text{for } v<V \tag{7}$$

$$PI_{iV-1} = \frac{1}{\psi}(r_V - \delta PI_{iV});\ \text{for } v=V \tag{8}$$

where Δ_v is a scalar factor that converts the continuous flow into stock increments and represents here the length of the time interval.

Equation (4) suggests that the marginal rate of substitution between consuming now and consuming later equals the relative price of consuming later instead of now times the intertemporal elasticity of substitution. Equations (7) and (8) indicate that the current price of investment depends on the rates of return to capital and depreciation adjusted to capital gain or loss (captured by the future price of investment). Equations (5), (6) and (7) are terminal conditions that insure that stocks of capital and foreign assets are constant in steady state.

---

[3] A reformulation of the continuous, infinite-horizon time optimization problem into a discrete, finite-horizon time problem requires dynamic aggregation methodology. See Lavoie *et al.* (2001)

(b) The Household Static Consumption/Investment Decision Problem

Domestic final demand decisions in each region i are made by a representative household. The decision problem of the household can be broken down to a "consumer" and "investor" problem.

The domestic consumer considers products of industries from different regions as imperfect substitutes [Armington (1969)]. This is represented by a two level utility function. The first level combines consumption goods $c_{.siv}$, assuming constant expenditure shares ρ_{si}. The second level determines the optimal composition of the consumption aggregate in terms of geographic origin.

Formally, the household's consumption preferences are given by a log-linear transformation of a Cobb-Douglas utility function at the first level, and a constant elasticity of substitution (CES) function at the second level:

$$U_{iv} = \sum_{s \in S} \rho_{si} \log c_{.siv} \quad \text{where} \quad \sum_{s \in S} \rho_{si} = 1$$

(9)

$$c_{.siv} = \left(\sum_{j \in W} \delta_{jsi} c_{jsiv}^{(\sigma_{si}-1)/\sigma_{si}} \right)^{\frac{\sigma_{si}}{(\sigma_{si}-1)}}$$

where:

c_{jsiv} is the consumption in region i of goods s produced in region j in period v,

$c_{.siv}$ is the composite of domestic and imported goods in period v,

δ_{jsi} are consumption share parameters in region i of goods s produced in region j, and

σ_{si} are the Armington elasticities of substitution for consumption in region i for good s.

The consumer maximizes (9) subject to

$$PC_{iv}C_{iv} = \sum_{j \in W} \sum_{s \in S} (1 + \tau_{jsiv}) p_{jsiv} c_{jsiv} \tag{10}$$

where p_{jsiv} denotes prices on which consumers have no influence, and the term on the left-hand side results from the intertemporal decision of the household.

The investor's problem, in turn, is to determine the optimal composition of the local investment good. The decision problem involves minimizing

$$Minimize \ PI_{iv}I_{iv} = \sum_{j \in W} \sum_{s \in S} (1 + \tau_{jsiv}) p_{jsiv} I_{jsiv} \tag{11}$$

Subject to

$$\log(I_{iv}) = \sum_{s \in S} \omega_{si} \log(I_{.siv}), \qquad \sum_{s \in S} \omega_{si} = 1, \tag{12}$$

$$I_{.siv} = \left\{ \sum_{j \in W} \delta_{jsi} I_{jsiv}^{\frac{\sigma_s - 1}{\sigma_s}} \right\}^{\frac{\sigma_s}{\sigma_s - 1}}$$

Firms

Each region is characterized by perfectly competitive industrial sectors. Demand for capital, labour and intermediate inputs by producers result from minimization of variable unit costs v_{isv}

$$v_{isv}Q_{isv} = \sum_{j \in W} \sum_{d \in S} (1 + \tau_{jdiv}) p_{jdiv} x_{jdisv} + w_{iv}L_{isv} + r_{iv}K_{isv} \tag{13}$$

subject to a Cobb Douglas production function

$$\log Q_{isv} = \alpha_l \ \log L_{isv} + \alpha_k \ \log K_{isv} + \sum_{d \in S} \alpha_{dis} \log x_{.disv} \tag{14}$$

157

where

$$x_{.disv} = \left(\sum_{j \in W} \beta_{jdis} x_{jdisv}^{(\sigma_{di}-1)/\sigma_{di}} \right)^{\frac{\sigma_{di}}{\sigma_{di}-1}} \tag{15}$$

are composite intermediate inputs in terms of geographical origin, x_{jdisv} is the amount of intermediate goods purchased by sector s of region i from sector d of region j in period v, and p_{jdiv} is the price of goods d sold by region j to region i in period v, and σ_{di} is the Armington elasticity of substitution of good d in region i.

To guarantee homogeneity of degree one of the unit costs in prices, we set

$$\alpha_{L_{is}} + \alpha_{K_{is}} + \sum_{d \in S} \alpha_{dis} = 1 \tag{16}$$

where α and β are share parameters and $\beta_{jdis} = 0, \forall j \neq i$ if d is non-tradable. Profit maximization, in this perfect competitive setting, implies prices equal marginal cost.

$$p_{isv} = v_{isv} \tag{17}$$

Instantaneous General Equilibrium

The instantaneous general equilibrium is defined as a static allocation, supported by a vector of prices ($p_{is.v}, P_{idjv}, w_{iv}, r_{iv}$), $i, j \in W$, consistent with the inter-temporal constraints and choices, such that
- Consumers maximize (9) subject to (10);
- Investors minimize (11) subject to (12);
- Firms minimize (13) subject to (14);
- Supply equals demand in each market:

158

$$Q_{isv} = \sum_{j \in W} \left[c_{isjv} + I_{isjv} + \sum_{d \in S} x_{isjdv} \right], \quad s \in S, \ i \in W; \qquad (18)$$

$$k_{iv} = \sum_{s \in S} k_{isv}, \quad i, j \in W$$

$$L_{iv}^{\text{sup}} = \sum_{s \in S} L_{isv}, \quad i \in W$$

where L_{iv}^{sup} denotes fixed labour supply.

Dataset and calibration procedure

The base year for the model's data set is 1999. The current model consists of five regions, three Canadian regions, the United States (USA), and the Rest of the World (ROW) aggregated as one region. The three Canadian regions are:

- Canada East comprising Atlantic Canada and Québec.
- Ontario
- Canada West comprising the Prairies, North West Territories, Nunavut, Alberta, British Columbia and the Yukon.

The forty-six commodities, level S, from the trade flow data were mapped into nine sectors. The sectoral mapping from the S level to the 24-sector breakdown of Version 1.0 and the 9-sector breakdown used in Version 2.0 are shown in Table 1.

Table 1: Sectoral Aggregation

9 Sector aggregation	24 Sector groups	S level groups
Agriculture and Forestry	Agriculture and Forestry	1,2,3,4
Food, Beverages & Tobacco	Food, Beverages & Tobacco	9,10,11,12
Textiles and Apparel	Textiles	14,15
Mining	Mining &Quarrying other than Petroleum	5,7,8
Wood and Related Products	Wood products	16,17,18,19
Other Manufacturing	Chemicals, Fertilizers & pharmaceuticals	6,13,20,21,25,26, 27,28
Machinery	Non-electrical machinery	22,24
Transport Equipment	Transport equipment	23
Services	Communication Services & Other Utilities	29,30,31,32,33,34, 35,36,38,39-46

Source: Statistics Canada, National Account Division, Table 386-0001, 1999

Data requirements for our model consist of nominal bilateral (international and inter-regional) trade flows; input-output tables, national accounts data (consumption demand by sector, labor and capital earnings).

Maintaining consistency among the sources is a challenging and time-consuming task. Therefore, many CGE models have turned to existing databases such as the Global Trade Assistance and Production (GTAP) data package. Despite the convenience, GTAP data had some major disadvantages: most importantly for our exercise, the GTAP database does not provide with Canadian provincial data.[4] For this reason, we have opted to develop our own database, collecting data from a variety of national and international sources.

The Canadian inter-provincial and international trade data were obtained from the National Accounts Division of Statistics Canada and Industry Canada Trade Data. The trade flows of the USA and the Rest of the World were retrieved from Industry Canada Trade Data (Strategis.ca), the World Trade Organizer and GTAP version 5.

The three Canadian economic regions were assumed to share the same production technology as Canada as a whole[5]. Therefore, the Canadian input-output table was used to derive the production technology coefficients; i.e., the share of intermediate inputs, labour and capital in final production. The Canadian input-output table was retrieved from CANSIM II database (tables 381009 and 3810010) for 1997. The Bureau of Economic Analysis provided the USA input and output tables, level M, for 1999.

Given the lack of an input-output matrix for the composite Rest of the World, intermediate technology coefficients were

---

[4] An additional consideration during the model building phase was the desire for more timely data; the GTAP database then available had data only for 1997 (GTAP version 5). The GTAP version 6, available since March 2005, updated the data to 2001.

[5] Due to confidentiality issues, provincial input-output tables have many cells with non-available data ("suppressed") that renders their use impractical.

generated on the basis of the GTAP database which provides estimates on intermediate demand by sector for the Rest of the World region. The value added for ROW, was retrieved from 'Sources of Factors Income', also from the GTAP 5 database.

The nest step after collecting the data is to ensure the consistency of the baseline data set. This requires that a) supply equals the demand for all commodities; b) all industries make no excess profits; c) all domestic agents' budget constraints are satisfied; and d) bilateral trade flows are mutually consistent.

The calibration of the static side of the model follows the approach developed by Srinivasan and Whalley (1986). The calibration of the dynamic side of the model follows the approach developed by Mercenier and Michel (1994), and Mercenier and Akitobi (1993). Table 2 reports on some parameters characterizing the dynamic behavior of the economy, assuming a finite-time horizon of 60 years.

Table 2: Some Parameters that Characterize the Dynamic Behaviour of Version 2.0

Parameter	Value
Number of periods	11
Time horizon	60 years
Discount rate ρ :	9.0 percent
Inter-temporal substitution elasticity $1/\gamma$:	1.0
Capital-output ratio (calibrated):	
Canada East:	4.65
Ontario	4.49
Canada West	5.06
USA	5.00
ROW	4.98
Depreciation rate of capital δ (%, calibrated):	
Canada East:	2.97
Ontario	3.10
Canada West	2.63
USA	3.01
ROW	2.67

Concluding Comments

The development of the dynamic CGE model, Version 2.0 described above, adds to the toolkit of International Trade Canada in evaluating quantitatively the implications of possible policies affecting Canada's trade and investment. It complements the static Version 1 model, the GTAP model which is also in use within the Department, as well as partial equilibrium and gravity models.

The present note describes the model to serve as a reference and as part of International Trade Canada's ongoing efforts to maintain transparency and to engage the professional trade policy and economics communities in its work.

References

Armington P. (1969), "A Theory of Demand for Products Distinguished by Place of Production", IMF Staff Papers, 16, pp159-178

Bhattacharyya, S.C. (1996) "Applied General Equilibrium Models for Energy Studies: A Survey", *Energy Economics*, 18, pp145-164

Burniaux, J.-M., Martin,J.P.,Nicoletti,G.,Martins,J.O.(1992) "Green-A Multi-Sector, Multi-Region General Equilibrium Model for Quantifying the Cost of Curbing CO_2 Emissions: A technical Manual", *Economics Department Working Paper No. 116*, Paris:OECD

GTAP (2001), Global Trade, Assistance, and Production: The GTAP 5 Data Package, Centre for Global Trade Analysis, Purdue University

GTAP-Dynam. Monash University, 2004. www.gtap.agecon.purdue.edu/models/dynamic.asp.

Klepper, G., S. Peterson, K. Springer (2003), "DART97: A Description of the Multi-regional. Multi-sectoral Trade Model for the Analysis of Climate Policies", *Kiel Working Paper No. 1149*, Kiel Institute for World Economics.

Lavoie, Claude, Mérette, M. and M. Souissi. (2001). "A Multi-sector Multi-country Dynamic General Equilibrium Model with Imperfect Competition", Department of Finance. Working Paper No. 2001-10.

Mercenier J., B.Akitoby,(1993), "On Intertemporal General-Equilibrium Reallocation Effects of Europe's Move to a Single Market", *Institute for Empirical Macroeconomics Working Paper 87*, Federal Reserve Bank of Minneapolis, Minneapolis,MN.

Mercenier J., P. Michel, (1994), "Discrete-time finite horizon approximation of infinite horizon optimization problems with steady-state invariance", *Econometrica*, Vol 62, No.3, pp635-656

Nordaus, W.D., Young Z. (1996), "A regional Dynamic General Equilibrium Model of Alternative Climate-Changes Strategies", American Economic Review, 86(4), pp741-765

Papadaki, E., M. Mérette, Y. Lan and J. Hernández, "The International Trade Canada Trade Model, Version 1.0", *Trade Policy Research 2004*, (Department of Foreign Affairs and International Trade: Ottawa, 2004): 247-264.

Pereira, A.M., Shoven, J.B. (1988), "Survey of dynamic computable general equilibrium models for tax policy evaluation", Journal of Policy Modeling, 10(3), pp401-436

Srinivasan T.N. and J. Whalley (eds.), 1986, *General Equilibrium Trade Policy Modeling*, The MIT Press.

Adapting the COMPAS model for Foreign Affairs and International Trade Canada

Shinji Kinjo[*]

Introduction and overview

The partial equilibrium model plays an important role in the quantitative analysis of trade and trade policy, complementing the computable general equilibrium model, which is the tool of choice for assessing the effects of major trade policy changes such as multilateral trade rounds or bilateral/regional free trade agreements, and the so-called "gravity model", which has proved to be particularly successful in helping understand the role of geography, both physical and human, in shaping global trade patterns. As Krugman and Obstfeld (2000:188) explain, trade policy for a particular sector or good can be reasonably well understood without going into detail about the ripple effects of that policy in the rest of the economy. Accordingly, the partial equilibrium model's relative simplicity and economy in terms of data requirements have made it the tool of choice for more narrowly focussed trade issues.

A particularly convenient set of partial equilibrium models was developed for the US International Trade Commission (USITC) by Francois and Hall (1997)[1], the Commercial Policy

[1] Joseph. F. Francois and H. Keith Hall (1997). "Partial Equilibrium Modeling" in Joseph F. Francois and Kenneth A. Reinert (eds.), *Applied Methods for Trade Policy Analysis — A Handbook* (U.S.A: Cambridge University Press): 122-155.

Analysis System (COMPAS) models. These models, although theoretically sound, require information on only a handful of key parameters and can be executed in standard spreadsheets. A good example of COMPAS models in practice was provided by the World Trade Organization's Article 22.6 arbitration regarding the economic and trade effects of the Continued Dumping and Subsidy Offset Act of 2000 (CDSOA), the well-known "Byrd Amendment". The United States submitted a number of COMPAS simulations to support its views regarding the size of the trade effects of the CDSOA disbursements.[2]

At the same time, the discussion of the reliability of the results of such simulations reveals the extent to which the results depend on the choice of key parameters.[3] In particular, the determination of the various elasticities that largely determine the outcome becomes one of the big and difficult problems in applying the partial equilibrium model.

This Chapter sets out the theoretical structure of the partial equilibrium model and develops a modified version of the COMPAS model for applications in International Trade Canada. The results of this model are compared to those derived using the COMPAS model by the United States in the context of the WTO arbitration on the trade effects of the Byrd Amendment disbursements..

The theoretical structure of the partial equilibrium model

Following Francois and Hall (2003), we describe the market that we want to analyze with a set of elasticities – i.e., elasticities of supply, aggregate elasticities of demand and elasticities of substitution. We concentrate on one market which has trade with two countries.[4] In order to identify each country, we use

---

[2] World Trade Organization (2004). *United States — Continued Dumping and Subsidy Offset Act of 2000 (WT/DS234/ARB/CAN)*. Retrieved January 20, 2005, from http://www.wto.org/

[3] WTO (2004); at p. 32.

[4] Francois and Hall (2003) deal with 25 markets.

the following suffixes: 1 = domestic country, 2 = exporting country A, and 3 = exporting country B.

The elasticities required for the model are defined as follows:

Elasticities of substitution in the domestic market

$$\sigma_{ij} = \sigma_{ji} = -\frac{d\ln(Q_i/Q_j)}{d\ln(P_i/P_j)} (>0) \tag{1}$$

where: σ_{ij}: elasticity of substitution ($i \neq j$)

Q_i: quantity of trade for goods made in country i (demand side)

P_i: price of goods made in country i (demand side)

Elasticities of supply to the domestic market

$$\varepsilon_{Si} = \frac{d\ln Q_i^S}{d\ln P_i^S} (>0) \tag{2}$$

where: ε_{Si}: elasticity of supply of goods made in country i.

Q_i^S: quantity of trade for goods made in country i. (supply side)

P_i^S: price of trade for goods made in country i. (supply side)

Aggregate elasticity of demand in the domestic market

$$N_A = \frac{d\ln Q_A}{d\ln P_A} (<0) \tag{3}$$

where : N_A: aggregate elasticity of demand.

Q_A: aggregate quantity of trade for goods. (demand side)

P_A: aggregate price of trade for goods. (demand side)

With these parameters, we can calibrate the own- and cross-price demand elasticities. Following Francois and Hall (1997:135-9), we derive these elasticities as follows:

Own-price elasticities of demand in the domestic market:

$$\eta_{ii} = \frac{\partial \ln Q_i}{\partial \ln P_i} = \theta_i N_A - \sum_{j \neq i} \theta_j \sigma_{ij} \cdots (< 0) \tag{4}$$

where θ_i : domestic market share (in quantity) of the good made in country i.

Cross-price elasticities of demand in the domestic market:

$$\eta_{ij} = \frac{\partial \ln Q_i}{\partial \ln P_j} = \theta_j (\sigma_{ij} + N_A) \cdots (> 0) \tag{5}$$

Using (2-4) and (2-5), the change in demand for the good made in country i in response to a change in its price (e.g., due to a tariff) can be depicted as follows:

$$d \ln Q_i = \eta_{ii} d \ln P_i + \eta_{ij} d \ln P_j + \cdots = \sum_j \eta_{ij} d \ln P_j \tag{6}$$

Using (2-2), we also get the change in supply of goods made in country i as follows:

$$d \ln Q_i^S = \varepsilon_{Si} d \ln P_i^S \tag{7}$$

Finally, a policy is added by introducing a tariff on goods i. The impact on the supply price is then given by the following equation:

$$d \ln P_i^S = \ln P_i^{S*} - \ln P_i^{S\#} = \ln(\frac{P_i^{S*}}{P_i^{S\#}}) \tag{8}$$

where: P_i^{S*} is the equilibrium price for goods made in country i (supply side)

$P_i^{S\#}$: offer price given the tariff (supply side)

The condition for market clearing can be written as follows:

$$d \ln Q_i = d \ln Q_i^S \tag{9}$$

$$P_i^* = P_i^{S*} \tag{10}$$

$$P_i = P_i^S \text{ except policy applied country.} \tag{10'}$$

where: P_i^* is the equilibrium price for goods made in country i
(demand side).

In other words, market behaviour is perfectly competitive; however domestic goods and imports are differentiated. In this case, we have three countries to consider. If we assume that the tariff is applied to the goods made in country 2, the equation can be written as follows:

$$\begin{cases} \eta_{11} \ln(\frac{P_1^*}{P_1}) + \eta_{12} \ln(\frac{P_2^*}{P_2}) + \eta_{13} \ln(\frac{P_3^*}{P_3}) = \varepsilon_{S1} \ln(\frac{P_1^*}{P_1}) \\ \eta_{21} \ln(\frac{P_1^*}{P_1}) + \eta_{22} \ln(\frac{P_2^*}{P_2}) + \eta_{23} \ln(\frac{P_3^*}{P_3}) = \varepsilon_{S2} \ln(\frac{P_2^*}{P_2^{S\#}}) \\ \eta_{31} \ln(\frac{P_1^*}{P_1}) + \eta_{32} \ln(\frac{P_2^*}{P_2}) + \eta_{33} \ln(\frac{P_3^*}{P_3}) = \varepsilon_{S3} \ln(\frac{P_3^*}{P_3}) \end{cases} \tag{11}$$

If we assume $\varepsilon_{Si} - \eta_{ii} \neq 0$, we obtain the solution as follows:

$$\begin{pmatrix} \ln(\frac{P_1^*}{P_1}) \\ \ln(\frac{P_2^*}{P_2}) \\ \ln(\frac{P_3^*}{P_3}) \end{pmatrix} = \begin{pmatrix} -1 & \frac{\eta_{12}}{\varepsilon_{S1} - \eta_{11}} & \frac{\eta_{13}}{\varepsilon_{S1} - \eta_{11}} \\ \frac{\eta_{21}}{\varepsilon_{S2} - \eta_{22}} & -1 & \frac{\eta_{23}}{\varepsilon_{S2} - \eta_{22}} \\ \frac{\eta_{31}}{\varepsilon_{S3} - \eta_{33}} & \frac{\eta_{32}}{\varepsilon_{S3} - \eta_{33}} & -1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ \frac{-\varepsilon_{S2}}{\varepsilon_{S2} - \eta_{22}} \ln(\frac{P_2^{S\#}}{P_2}) \\ 0 \end{pmatrix}$$

$$\tag{12}$$

The equation above assumed that the determinant of the matrix is not zero.

$$\det A = \begin{vmatrix} -1 & \dfrac{\eta_{12}}{\varepsilon_{S1} - \eta_{11}} & \dfrac{\eta_{13}}{\varepsilon_{S1} - \eta_{11}} \\[3mm] \dfrac{\eta_{21}}{\varepsilon_{S2} - \eta_{22}} & -1 & \dfrac{\eta_{23}}{\varepsilon_{S2} - \eta_{22}} \\[3mm] \dfrac{\eta_{31}}{\varepsilon_{S3} - \eta_{33}} & \dfrac{\eta_{32}}{\varepsilon_{S3} - \eta_{33}} & -1 \end{vmatrix} \neq 0$$

The possibility that either $\varepsilon_{Si} - \eta_{ii} = 0$ or that $\det A = 0$ cannot be excluded although these are unlikely outcomes. In the event that these values did obtain, some modifications to the above analysis would be necessary.

Validating the Model: reproducing COMPAS model results

Various versions of the COMPAS model are available from Francois and Hall (1998), including:
1) Price Dumping Model,
2) Cost Dumping Model,
3) Subsidy Model, and
4) Tariff Model.

To demonstrate the consistency of our model with the COMPAS model, we consider first the price dumping model.

The definition of a dumping margin is as follows (Boltuck et al. 1991:155):

$$DM = \frac{FV - P}{P}$$

where: DM is the dumping margin (%),
FV is "fair value" (price without dumping)
P is dumped price.

Therefore with dumping distortion, we get the price as follows:

$$P = \frac{FV}{1+DM}$$

Without consideration of domestic tariff rate or transportation costs, we can introduce this distortion directly as follows:

$$P_2^{S\#} = \frac{P_2}{1+DM}$$

$$
\begin{pmatrix} \ln(\frac{P_1^*}{P_1}) \\ \ln(\frac{P_2^*}{P_2}) \\ \ln(\frac{P_3^*}{P_3}) \end{pmatrix} = \begin{pmatrix} -1 & \frac{\eta_{12}}{\varepsilon_{S1}-\eta_{11}} & \frac{\eta_{13}}{\varepsilon_{S1}-\eta_{11}} \\ \frac{\eta_{21}}{\varepsilon_{S2}-\eta_{22}} & -1 & \frac{\eta_{23}}{\varepsilon_{S2}-\eta_{22}} \\ \frac{\eta_{31}}{\varepsilon_{S3}-\eta_{33}} & \frac{\eta_{32}}{\varepsilon_{S3}-\eta_{33}} & -1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ \frac{-\varepsilon_{S2}}{\varepsilon_{S2}-\eta_{22}}\ln(\frac{P_2^{S\#}}{P_2}) \\ 0 \end{pmatrix}
$$

From this equation, as shown in Appendix Table 1, we obtain the same results as Francois and Hall (1998).

To demonstrate consistency with the COMPAS tariff model, we introduce tariffs into our model. If we put the tariff in country 2 at the rate t %, then the price of imports from country 2 at the domestic country can be written as follows:

$$P_2^{S\#} = (1+t)P_2$$

Francois and Hall (1998) deal with two countries in their tariff model; accordingly, we derive the equation as follows. The consistency of results with the COMPAS model is demonstrated in Appendix Table 2

$$
\begin{pmatrix} \ln(\frac{P_1^*}{P_1}) \\ \ln(\frac{P_2^*}{P_2}) \end{pmatrix} = \begin{pmatrix} -1 & \frac{\eta_{12}}{\varepsilon_{S1}-\eta_{11}} \\ \frac{\eta_{21}}{\varepsilon_{S2}-\eta_{22}} & -1 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ \frac{-\varepsilon_{S2}}{\varepsilon_{S2}-\eta_{22}}\ln(\frac{P_2^{S\#}}{P_2}) \end{pmatrix}
$$

Validating the model: comparison to COMPAS model results developed for the Byrd Amendment arbitration.

On November 11, 2004, seven countries (Canada, Japan, EU, Mexico, Brazil, India and Korea—the "Requesting Parties")[5] requested authorization from the WTO to suspend concessions to the United States for its failure to comply with the WTO's recommendation that the United States withdraw the Continuing Dumping and Subsidy Offset Act (CDSOA). In the course of the arbitration, the United States provided analysis describing the trade effects of the CDSOA disbursement in various goods sectors (WTO 2004:28).

For the purposes here, we consider the analysis of the effect of the CDSOA disbursement on Canadian alloy magnesium exports to the US market. In this case, the trade policy measure in question is equivalent to a domestic subsidy.

For this application, the partial equilibrium equation (12) can be written as follows:

$$
\begin{pmatrix} \ln(\dfrac{P_1^*}{P_1}) \\ \ln(\dfrac{P_2^*}{P_2}) \\ \ln(\dfrac{P_3^*}{P_3}) \end{pmatrix} = \begin{pmatrix} -1 & \dfrac{\eta_{12}}{\varepsilon_{S1} - \eta_{11}} & \dfrac{\eta_{13}}{\varepsilon_{S1} - \eta_{11}} \\ \dfrac{\eta_{21}}{\varepsilon_{S2} - \eta_{22}} & -1 & \dfrac{\eta_{23}}{\varepsilon_{S2} - \eta_{22}} \\ \dfrac{\eta_{31}}{\varepsilon_{S3} - \eta_{33}} & \dfrac{\eta_{32}}{\varepsilon_{S3} - \eta_{33}} & -1 \end{pmatrix}^{-1} \begin{pmatrix} \dfrac{-\varepsilon_{S1}}{\varepsilon_{S1} - \eta_{11}} \ln(\dfrac{P_1^{S\#}}{P_1}) \\ 0 \\ 0 \end{pmatrix}
$$

$$(13)$$

The suffix identifiers are in this application are as follows:

1   = domestic country (US);
2   = exporting country (Canada); and
3   = exporting country (Rest of the World).

[5] Chile joined the group December 7, 2004.

The domestic subsidies, by reducing costs for US domestic industry result in a price change that can be written as follows:

$$P_1^{S\#} = \frac{C}{S + C} \tag{14}$$

where: C represents the costs for production of the US domestic industry without subsidy; and

S is the subsidy paid by the government

In the Canadian 2001 case, C is assumed to be equal to the value of domestic shipments (US$158 million), while S was equal to US$687,716. So the price reduction can be calculated as being about 0.435%. The model parameters provided by the United States were as follows:

Key parameters, Alloy Magnesium from Canada, 2001	
elasticity of substitution	5
elasticity of domestic supply to US market	7.5
elasticity of Canadian import supply to US market	100
elasticity of other countries import supply to US market	10
aggregate elasticity of demand	-0.875

As we can see in Appendix Table 3, the trade effect produced by the model with the above assumptions is a reduction of Canadian exports by US$ 342,357.

It is of interest to compare the results of the above calculation with the model submitted by the Requesting Parties to the arbitrator (WTO 2004:27).

$$\Delta V_m = \left(\frac{\Delta Q_m / Q_m}{\Delta P_d / P_d} \right) * (\Delta P_d * Q_d) * \left(\frac{P_m Q_m}{P_d Q_d} \right) \tag{15}$$

where: ΔV_m is the reduction in value of imports,

ΔQ_m : reduction in quantity of imports,

Q_m : quantity of imports,

ΔP_d : reduction in domestic shipment price,

P_d : domestic shipment price,

Q_d : quantity of domestic shipment,

P_m : import price.

The arbitrator determined that the Requesting Parties' model should be modified by adding a term to capture the less-than-100 percent pass through of the CDSOA disbursements into the offer price of US suppliers (WTO 2004:33). Taking into account the pass-through effect, equation (15) can be written as follows:

$$\Delta V_m = \eta * \alpha S * R \tag{16}$$

where: $\eta = \dfrac{\Delta Q_m / Q_m}{\Delta P_d / P_d}$: the elasticity of substitution,

α is the rate of pass through of the subsidy to prices

$S = \Delta P_d * Q_d$: the total value of payments,

$R = \dfrac{P_m Q_m}{P_d Q_d}$: the ratio of the value of imports to the

value of domestic shipments in the markets in question.

At a glance, the modified Requesting Parties model (16) appears to be quite different from the partial equilibrium model developed earlier or the COMPAS model on which it is based. What is the difference between them? In order to reconcile these models, we use our model (13) and reduce the number of countries from three to two. Thus, we rewrite (13) as follows,

$$
\begin{pmatrix} \ln(\dfrac{P_1^*}{P_1}) \\ \ln(\dfrac{P_2^*}{P_2}) \end{pmatrix} =
\begin{pmatrix} -1 & \dfrac{\eta_{12}}{\varepsilon_{S1} - \eta_{11}} \\ \dfrac{\eta_{21}}{\varepsilon_{S2} - \eta_{22}} & -1 \end{pmatrix}^{-1}
\begin{pmatrix} \dfrac{-\varepsilon_{S1}}{\varepsilon_{S1} - \eta_{11}} \ln(\dfrac{P_1^{S\#}}{P_1}) \\ 0 \end{pmatrix}
$$

Assuming the initial prices to be unitary, we derive the change in the value of imports as follows:

$$\ln(\dfrac{V_2^*}{V_2}) = -\dfrac{(\varepsilon_{S2}+1)\varepsilon_{S1}\eta_{21}}{\eta_{12}\eta_{21} - (\varepsilon_{S1} - \eta_{11})(\varepsilon_{S2} - \eta_{22})} \ln(P_1^{S\#}) \tag{17}$$

Assuming perfect competition in the initial stage, the cost of production of domestic producer is equal to the value of domestic shipments. So, with using (14) we can rewrite (17) as follows:

$$\ln(\frac{V_2^*}{V_2}) = \eta_{agg} \ln(\frac{V_1 + \alpha S}{V_1}) \tag{18}$$

where $\eta_{agg} = \dfrac{(\varepsilon_{S2} + 1)\varepsilon_{S1}\eta_{21}}{\eta_{12}\eta_{21} - (\varepsilon_{S1} - \eta_{11})(\varepsilon_{S2} - \eta_{22})}$

Finally, we apply the Taylor expansion to (18) to derive the following expression:

$$\Delta V_2 = \eta_{agg} * \frac{\alpha S}{V_1} * V_2 + \eta_{agg} * (\eta_{agg} - 1) * \frac{(\frac{\alpha S}{V_1})^2}{2!} * V_2 + \cdots \tag{19}$$

Comparing with (16) and (19), and taking into account the fact that $\eta = \eta_{agg}$ and $R = \dfrac{V_2}{V_1}$,

we can conclude that the Requesting Parties' model (3-4) is in fact the first order approximation of our model or COMPAS.

Conclusion

As this Chapter shows, the partial equilibrium model developed comes from the same theoretical background as does the COMPAS model. The difference is the way the equations are solved. COMPAS uses Cramer's Rule; the approach taken here relies on the inverse matrix directly.

As a trial empirical analysis, the results are validated by reproducing the COMPAS-based estimates of trade effects calculated by the United States in the Byrd Amendment arbitration. As well, the relationship between the partial equilibrium model developed here, based on COMPAS, and the model used by the arbitrator in Byrd Amendment is set out.

The theoretical framework for the partial equilibrium model is straightforward. It is a simple matter to add additional countries for analytical purposes. And the model can be edited freely. That being said, the difficulties of applying the model in novel situations where the parameters of the model must be estimated from actual data should not be underestimated. However, as the United States stated (WTO 2004:25), "The fact that an exercise is complex, tedious, or even burdensome does not mean it can be dispend with." As practitioners in the trade policy field, we officers should brush up our literacy about economic models. I wish this paper can make a contribution for that purpose.

Bibliography

Boltuck, Richard., Francois, Joseph F. and Kaplan, Seth. (1991). The Economic Implications of the Administration of the U.S. Unfair Trade Laws. In Boltuck, Richard and Litan, Robert E. (Ed.), Down in the Dumps – Administration of the Unfair Trade Laws (pp152-99). Washington, D.C.: The Brookings Institution.

Francois, Joseph and Hall, H. Keith. (2003). Global Simulation Analysis of Industry-Level Trade Policy. Retrieved February 3, 2005, from http://www.intereconomics.com/handbook/Models/Spreadsh eet Models/GSIMpaper.pdf

Francois, Joseph and Hall, Keith. (1998). COMPAS Commercial Policy Analysis System. Retrieved February 8, 2005, from http://www.intereconomics.com/handbook/Models/Spreadsh eet%20Models/COMPAS1.XLS

Francois, Joseph. F. and Hall, H. Keith. (1997). Partial Equilibrium Modeling. In Francois, Joseph F. and Reinert, Kenneth A (eds.), Applied Methods for Trade Policy Analysis – A Handbook(pp122-55). U.S.A: Cambridge University Press.

Gallaway, Michael P. McDaniel, Christine A. and Rivera, Sandra A. (2003). Short-run and long-run industry-level estimates of U.S. Armington elasticities. The north American Journal of Economics and Finance, Vol. 14, 49-68.

Krugman, Paul R. and Obstfeld, Maurice. (2000). International Economics Theory and Policy 5th ed. U.S.A: Addison-Wesley.

United States International Trade Commission (US-ITC). (2003). Industry & Trade Summary - pasta (USITC Publication 3592). Retrieved May 2, 2005, from http://prototype.usitc.gov/

United States International Trade Commission (US-ITC). (2002). Year in Review (USITC Publication 3594). Retrieved May 2, 2005, from http://prototype.usitc.gov/

United States International Trade Commission (US-ITC). (1999). Shifts in U.S. Merchandise Trade in 1998 (USITC Publication 3220). Retrieved May 2, 2005, from http://prototype.usitc.gov/

World Trade Organization (WTO). (2004). United States – Continued Dumping and Subsidy Offset Act of 2000(WT/DS234/ARB/CAN). Retrieved January 20, 2005, from http://www.wto.org/

Appendix Table 1: Price Dumping Model (reproducing COMPAS model result #1)

Partial Equilibrium Model for Domestic (country l)
Inputs
Output

Inputs (Elasticities)

Country	Elasticities of substitution		
	1	2	3
1	-	**3**	**3**
2	3	-	**3**
3	3	3	-

Country	Elasticities of supply to demand market		
	1	2	3
	5	**10**	**10**

Aggregate Elasticities of demand			**-0.5**

Inputs (Market Data - initial state)

Country	Import Price (F.O.B)	Transaction Cost	MFN	Market Price
1	N.A.	N.A.	N.A.	1
2	**1**	**0%**	**0%**	1
3	**1**	**0%**	**0%**	1

Country	Quantity	Share of Quantity	Market Value (F.O.B)	Share of Value
1	113,142,857	55%	$113,142,857	55%
2	72,000,000	35%	$72,000,000	**35%**
3	20,571,429	10%	$20,571,429	**10%**
Sum	205,714,286	100%	$205,714,286	100%

Inputs (Market Distortion for supply side - offering state)

Country	Price Dumping Rate	Transaction Cost	Cost Dumping (NG)	Subsidy
1	**0%**	**0%**	**0%**	**0%**
2	20%	30%	0%	0%
3	0%	30%	0%	0%

Country	Tariff(Tax)	Price Change
1	**0%**	100.00%
2	**0%**	86.67%
3	**0%**	100.00%

Calibration

Own elasticities of domestic demand

Country	1	2	3	Own elasticities
1	-0.275	1.050	0.300	-1.625
2	1.650	-0.175	0.300	-2.125
3	1.650	1.050	-0.050	-2.750

Cross elasticities of domestic demand

Country	1	2	3
1	-	0.875	0.250
2	1.375	-	0.250
3	1.375	0.875	-

Equation

From Demand change

-1.000	0.132	0.038
0.113	-1.000	0.021
0.108	0.069	-1.000

Inverse Matrix

-1.020	-0.138	-0.041
-0.118	-1.017	-0.025
-0.118	-0.085	-1.006

From Supply change

0.000
0.118
0.000

Solution

$Ln(P1*/P1)$	-1.6%
$Ln(P2*/P2)$	-12.0%
$Ln(P3*/P3)$	-1.0%
$Ln(Q1*/Q1)$	-8.1%
$Ln(Q2*/Q2)$	23.0%
$Ln(Q3*/Q3)$	-10.0%
$Ln(V1*/V1)$	-9.7%
$Ln(V2*/V2)$	11.0%
$Ln(V3*/V3)$	-11.0%

Percentage Change (Same as estimated impacts in Francois and Hall 2003)

P1	-1.6%
P2	-11.3%
P3	-1.0%
Q1	-7.8%
Q2	25.9%
Q3	-9.5%
V1	-9.3%
V2	11.7%
V3	-10.4%

Appendix Table 2: Tariff Model
(reproducing COMPAS model result #2)

Partial Equilibrium Model for Domestic (country l)
Inputs
Output

Inputs (Elasticities)

	Elasticities of substitution		
Country	1	2	3
1	-	3	3
2	3	-	3
3	3	3	-
	Elasticities of supply to demand market		
Country	1	2	3
	5	10	10
Aggregate Elasticities of demand			-0.5

Inputs (Market Data - initial state)

Country	Import Price (F.O.B)	Transaction Cost	MFN	Market Price
1	N.A.	N.A.	N.A.	1
2	1	0%	0%	1
3	1	0%	0%	1

Country	Quantity	Share of Quantity	Market Value (F.O.B)	Share of Value
1	1,000,000,000	50%	$1,000,000,000	50%
2	1,000,000,000	50%	$1,000,000,000	50%
3	-	0%	-	0%
Sum	2,000,000,000	100%	$2,000,000,000	100%

Inputs (Market Distortion for supply side - offering state)

Country	Price Dumping Rate	Transaction Cost	Cost Dumping (NG)	Subsidy
1	0%	0%	0%	0%
2	0%	0%	0%	0%
3	0%	0%	0%	0%

Country	Tariff(Tax)	Price Change
1	0%	100%
2	10%	110%
3	10%	110%

182

Calibration

Own elasticities of domestic demand

Country	1	2	3	Own elasticities
1	-0.25	1.50	0	-1.75
2	1.50	-0.25	0	-1.75
3	1.50	1.50	0	-3

Cross elasticities of domestic demand

Country	1	2	3
1	-	1.25	0
2	1.25	-	0
3	1.25	1.25	-

Equation

From Demand change

-1.000	0.185	0
0.106	-1.000	0
0.096	0.069	-1.000

Inverse Matrix

-1.020	-0.189	0
-0.109	-1.020	0
-0.109	-0.116	-1

From Supply change

0
-0.081
-0.073

Solution

$Ln(P1*/P1)$	1.5%
$Ln(P2*/P2)$	8.3%
$Ln(P3*/P3)$	8.3%
$Ln(Q1*/Q1)$	7.7%
$Ln(Q2*/Q2)$	-12.6%
$Ln(Q3*/Q3)$	-12.6%
$Ln(V1*/V1)$	9.2%
$Ln(V2*/V2)$	-4.3%
$Ln(V3*/V3)$	-4.3%

Percentage Change (Same as estimated impacts in Francois and Hall 2003)

P1	1.5%
P2	8.6%
P3	8.6%
Q1	8.0%
Q2	-11.8%
Q3	-11.8%
V1	9.6%
V2	-4.2%
V3	-4.2%

Appendix Table 3: Model for the Byrd Amendment arbitration (reproducing COMPAS model result #3)

Partial Equilibrium Model for Domestic (country l)
Inputs
Output

Inputs (Elasticities)

Country	Elasticities of substitution		
	1	2	3
1	-	5	5
2	5	-	5
3	5	5	-

Country	Elasticities of supply to demand market		
	1	2	3
	7.5	100	10
Aggregate Elasticities of demand		-0.875	

Inputs (Market Data - initial state)

Country	Import Price (F.O.B)	Transaction Cost	MFN	Market Price
1	N.A.	N.A.	N.A.	1
2	1	0%	0%	1
3	1	0%	0%	1

Country	Quantity	Share of Quantity	Market Value (F.O.B)	Share of Value
1	158,000,000	54%	$158,000,000	54%
2	44.769,000	15%	$44.769,000 (b)	15%
3	88,202,000	30%	$88,202,000	30%
Sum	290,971,000	100%	$290,971,000	100%

Inputs (Market Distortion for supply side - offering state)

Country	Price Dumping Rate	Transaction Cost	Cost Dumping (NG)	Subsidy
1	0%	0%	0%	.435%
2	0%	0%	0%	0%
3	0%	0%	0%	0%

Country	Tariff(Tax)	Price Change
1	0%	99.57%
2	0%	100%
3	0%	100%

Calibration

Own elasticities of domestic demand

Country	1	2	3	Own elasticities
1	-0.475	.769	1.516	-2.760
2	2.715	-0.135	1.516	-4.365
3	2.715	.769	-.265	-3.7495

Cross elasticities of domestic demand

Country	1	2	3
1	-	0.635	1.250
2	2.240	-	1.250
3	2.240	0.635	-

Equation

From Demand change

-1.000	0.062	0.122
0.021	-1.000	0.012
0.163	0.046	-1.000

Inverse Matrix

-1.022	-0.069	-0.125
-0.024	-1.002	-1.015
-0.168	-0.058	-1.021

From Supply change

0.003
0
0

Solution

Ln(P1*/P1)	0.3%
Ln(P2*/P2)	0.0%
Ln(P3*/P3)	-0.1%
Ln(Q1*/Q1)	0.8%
Ln(Q2*/Q2)	-0.8%
Ln(Q3*/Q3)	-0.5%
Ln(V1*/V1)	0.5%
Ln(V2*/V2)	-0.8%
Ln(V3*/V3)	-0.6%

Percentage Change (Same as estimated impacts in Francois and Hall 2003)

P1	**-0.3%**
P2	**0.0%**
P3	**-0.1%**
Q1	**0.8%**
Q2	**-0.8%**
Q3	**-0.5%**
V1	**0.5%**
V2	**-0.8%**
V3	**-0.6%**

After Distortion (Trade Values etc.)

P1*	0.9968
P2*	0.9999
P3*	0.9995
Q1*	159,307,491
Q2*	44,430,020
Q3*	87,733,977
V1*	158,791,477
V2*	44,426,643 (b)
V3*	87,687,311

Memo: Reduction = (a) – (b) = $342,357

Trade Specialization in the Gravity Model of International Trade

Dan Ciuriak and Shinji Kinjo[*]

Introduction

The gravity model of international trade has become one of the standard tools for analyzing trade patterns and trade. In its usual modern articulation, the gravity model hypothesizes that the larger, the richer, the closer together two countries are, the more they trade. And they trade more intensively the more things they have in common, such as currency, language, shared political histories or colonial connections, a border, etc.. Coastal states trade more than landlocked states because they connect more easily. Standard gravity models explain about two-thirds of the variation in global trade, leaving only one-third to be explained by other trade theories.

One criticism of the gravity model is that it takes no account of comparative advantage, which still forms the bedrock of economists' understanding of international trade. This critique is particularly important when the gravity model is considered for policy applications such as identifying priority markets for trade promotion programs. For example, the potential for trade expansion through programs to reduce the impact of various types of "frictions" that impede trade might be greater with countries that have complementary patterns of comparative

advantage, as opposed to with countries with similar patterns of comparative advantage.

To address this problem, we introduce a trade specialization variable into a gravity model to capture the degree of complementarity of trading partners' comparative advantage. This short paper describes the test of this thesis.

Our general findings are as follows. The trade specialization index clearly distinguishes countries that are generally believed to be "most similar" from those that are believed to "most different". Its explanatory power in the gravity equation is good, comparing well with other established variables and it improves the overall goodness of fit of the gravity equation. It thus appears to be a useful addition to the gravity model toolkit. At the same time, it still leaves a large residual unexplained variation that weakens the gravity model's utility for some policy applications.

The balance of this note is organized as follows: the next section describes the construction of the trade specialization index that captures the degree of complementarity of two trading partner's comparative advantage, and of the gravity model developed to test this variable. The third section comments on the results of the estimation. The final section provides some concluding comments.

Capturing comparative advantage in a gravity model

From a practical perspective in applying the gravity model, it is of interest to take explicit account of comparative advantage. A well-established empirical approach to describing a country's comparative advantage is to infer it from the commodity structure of the country's trade. A variety of indexes have been established for this purpose; for our purposes, the most intuitively appealing measure is the Trade Specialization Index (TSI). This is defined as net exports (exports minus imports) in a given sector divided by total two-way trade in that sector. The range of this variable is from 1 if a country only exports in a given sector to −1 if it only imports in that sector.

We calculate a country's TSI for each sector defined at the 2-digit HS code level. We then calculate the simple correlation coefficient between two trading partner's TSIs. This variable can take values from 1 if the TSIs are identically distributed over the various sectors, to −1 if the TSIs are perfectly negatively correlated. Pairs that have an TSI correlation in positive territory would tend to be natural competitors in international trade while those in negative territory would tend to be natural trading partners, according to the principle of comparative advantage.

The pattern of TSI correlation corresponds well to general notions of similarity and dissimilarity of economies. For Canada, the most similar country is Australia with an TSI correlation of 0.57. A highly dissimilar country is China with an TSI correlation of -.39.

Table 1
Trade Specialization Correlation: Canada an its trading partners

Partner Country	RCA2004
Australia	0.57
Russia	0.51
Argentina	0.44
Norway	0.42
Finland	0.41
Chile	0.40
South Africa	0.40
Iceland	0.34
Venezuela	0.31
Sweden	0.31
United States	0.29
New Zealand	0.29
Denmark	0.23
Luxembourg	0.22
Peru	0.17
Brazil	0.15
Poland	0.15
Greece	0.14

Netherlands	0.14
France	0.12
United Kingdom	0.11
Colombia	0.10
Germany	0.06
Austria	0.06
Ireland	0.05
Singapore	0.03
Belgium	0.01
Spain	0.00
Turkey	-0.04
Korea	-0.05
Sri Lanka	-0.08
Portugal	-0.12
Mexico	-0.14
Malaysia	-0.15
Japan	-0.15
Philippines	-0.15
Thailand	-0.19
Switzerland	-0.20
India	-0.24
Indonesia	-0.24
Hong Kong	-0.29
Taiwan	-0.31
China	-0.39
Italy	-0.40

To test the explanatory power of this variable, we apply a standard gravity equation to a reduced set of countries, choosing the 44 largest trading economies (with the exception of Taiwan, for which the full set of gravity model data was not readily available). The gravity equation that we estimate is as follows:

$$\ln F_{ij} = const. + \ln D_{ij} + ComLang_{ij} + Landlocked_i + Landlocked_j + RCA_{ij}$$
$$+ RTA01_{ij} + RTA02_{ij} + RTA03_{ij} + \ln GDP_i + \ln GDP_j + \ln Area_i + \ln Area_j + \varepsilon_{ij}$$

(eq.1)

where

F_{ij} : amount of trade (export + import) between i and j.

D_{ij} : distance from i to j.

ComLang: Common Language (when at least 9% of popul-
tion of both countries uses same language, then unit, other-
wise the variable takes the value zero.)

Landlocked: If country i or j has no coast, then unit, otherwise zero.

RCA: the correlation of a trading pair's trade specialization indexes, as a measure
of the degree of complementary of their respective revealed comparative advan-
tage.

GDP: GDP (at current price in USD millions)

RTA_{ij} : Regional Trade Agreement between countries i and j

RTA01: Free Trade Agreement and Service Agreement with
EFTA.

RTA02: ASEAN Free Trade Area

RTA03: General System of Trade Preferences among de-
veloping countries.

$Area_i$: area of country i or j

The sources for the data used to estimate the model are set out
in Table 2.

Table 2
Sources of data

Variable	Unit	Source
F	million USD	World Trade Atlas
D	km	cepii
comlang_ethno	dummy	cepii
Landlocked	dummy	cepii
Area	km2	cepii
GDP	million USD	UN
RCA2002	-	Ciuriak and Kinjo (2005)*
RTA0x	dummy	WTO
* based on trade data from World Trade Atlas		

Results

Using 2002 data, we obtain the following regression result (see Table 3).

Table 3
The result of estimation

Variables	Coefficient	t-value
Log Distance	-0.80	-27.77
Common Language	0.79	10.32
Landlocked Importer	-0.47	-4.43
Landlocked Exporter	-0.69	-6.56
RCA	-0.49	-4.13
RTA01	-1.85	-4.26
RTA02	1.70	6.55
RTA03	0.48	5.60
Log Exporter GDP	0.95	46.99
Log Importer GDP	1.00	49.52
Log Area of Importer	-0.08	-6.49
Log Area of Exporter	-0.13	-10.57
Summary statistics		
Constant	-34.89	
Number of observations	1880	
R^2	0.80	
Root MSE	1.058	

The coefficients of distance and GDP have the usual high level of significance found in other gravity models. Other variables have the right sign and, as typical of gravity models, a good level of significance. As we can see, the variable constructed to capture the complementarity of a pair of trading partner's trade patterns (RCA) has a reasonably good level of significance, as well. Overall, the model fits very well, with an R^2 of .80.

Nonetheless, the RCA variable reduces the amount of unexplained variation by less than might be hoped for, given the expected importance of comparative advantage in shaping global trade flows. Table 4 highlights this; it compares the estimated level of two-way trade between Canada and each partner country with actual trade. The difference ("over-trading" if actual trade exceeds estimated trade, and "under-trading" if the reverse holds) can be contrasted with the pattern of RCA complementarity or similarity. Trading partners with complementary patterns of comparative advantage are shaded. As can be seen, most of Canada's trading partners with complementary trade patterns "over-trade" and are clustered at the top of the table; they account for the most intense degree of over-trading and the bulk of the volume. Conversely, those countries that "under-trade" tend to have similar trade patterns. This suggests that perhaps more of the variation could be explained by comparative advantage in a more refined model.

Table 4
Comparison between the result of estimation and actual trade volume in 2002, millions of US dollars or percent

Trade partner	Estimated level of two-way trade	Actual level of two-way trade (a)	Over- or under-trading (b)	Percentage over- or under-trading (b)/(a)
China	3,194	12,828	9,635	75%
Mexico	4,357	9,654	5,297	55%
Japan	11,544	15,148	3,604	24%
Norway	780	3,124	2,344	75%
Malaysia	250	1,600	1,349	84%
Thailand	363	1,470	1,107	75%

Korea	3,496	4,385	889	20%
Australia	1,181	1,841	660	36%
Venezuela	544	1,132	588	52%
Indonesia	378	928	550	59%
Brazil	1,256	1,702	446	26%
Sweden	939	1,367	428	31%
Chile	200	606	406	67%
Philippines	517	907	390	43%
New Zealand	297	489	191	39%
Austria	620	796	177	22%
Singapore	829	996	167	17%
Finland	493	633	140	22%
Peru	233	294	61	21%
Denmark	884	912	28	3%
Sri Lanka	56	84	28	33%
South Africa	458	473	15	3%
Luxembourg	94	108	13	12%
Iceland	62	69	7	11%
Colombia	474	469	-4	-1%
Argentina	259	234	-26	-11%
Netherlands	2,199	2,067	-133	-6%
Turkey	656	427	-229	-54%
Ireland	1,686	1,340	-346	-26%
Russia	775	394	-381	-97%
Greece	542	161	-381	-237%
Hong Kong	1,828	1,403	-424	-30%
Poland	742	285	-458	-161%
Portugal	736	251	-485	-193%
Switzerland	2,258	1,295	-963	-74%
Germany	8,188	7,166	-1,022	-14%
Belgium	3,161	2,109	-1,052	-50%
India	2,914	1,276	-1,638	-128%
Spain	2,990	1,258	-1,732	-138%
Italy	5,803	3,777	-2,026	-54%
UK	15,591	9,027	-6,564	-73%
France	12,985	5,013	-7,972	-159%
United States	421,008	359,017	-61,991	-17%
Total	517,822	458,514	-59,308	-13%

Two things stand out from this table apart from the pattern of over- and under-trading compared to the model's predictions: the highest volume of under-trading is with the three countries with which Canada has had the closest historical relationship, the UK, France and the United States. Meanwhile, China, which puts the gleam in the eye of those looking for greater diversification of Canada's trade already accounts for a much larger share of Canada's trade than would be expected given the kinds of factors that gravity models take into account..

Conclusion

The gravity model is one of the most empirically successful models in the economic toolkit. It serves to emphasize the importance of geography, both physical and human, in shaping the pattern of global economic interaction. Yet, by ignoring the most fundamental theory of trade, comparative advantage, it meets with skepticism in policy applications.

We have developed a way to augment the standard gravity model to take explicit account of comparative advantage using the correlation of trading partners' revealed comparative advantage. The results are intuitively plausible and statistically reasonably robust.

We believe this to be a practically useful development of the gravity model although admittedly one would have hoped for a greater reduction in the amount of unexplained variation in trade patterns from integrating comparative advantage into the gravity model.

The Role of Embassies and Consulates in Promoting Trade: Does Economic Freedom in the Host Economy Matter?

Dan Ciuriak and Shinji Kinjo[*]

Introduction

Are exports systematically associated with diplomatic representation abroad? Given the expenses of maintaining posts abroad, this question is obviously of deep interest to governments seeking to promote their country's export performance while coping with budgetary pressures.[1] In a recent article, Andrew Rose found that, yes, posts abroad do systematically increase trade with each post adding 6-10 percent additional exports; the effect is non-linear, with the first post adding more and additional posts adding less, suggesting diminishing returns to representation.[2]

To be sure, Rose's findings represent only the opening salvo in what is undoubtedly going to be a hotly debated issue. There are many unanswered questions about how exactly these export promotion gains are made (which goes to the microeconomic question of what is the role of government in this activity). Moreover, the model used to generate these results, the so-called "gravity model" of inter-

---

[1] For example, Canada's recent International Policy Statement set out ambitious targets for export growth in key emerging markets. How best to achieve these goals depends on the cost-effectiveness of alternative instruments of export promotion, one of which is the role of posts abroad.

[2] Andrew K. Rose, "The Foreign Service and Foreign Trade: Embassies as Export Promotion", *NBER Working Paper* No. 11111, February 2005.

national trade, although gaining increasing acceptance in theoretical terms, is only starting to gain standing in policy debates. A lot of people will need convincing before strategic decisions are made on the basis of this model's predictions.

Yet, the success of the gravity model in explaining the general pattern of global trade cannot be ignored and the desire to ground policy decisions in empirical research is intensifying. As an additional tool in a trade ministry's toolkit, its value needs to be better understood.

In this spirit, we take up a question that immediately arises with Rose's thesis: does it matter whether the host economy is economically "free", in the sense that the host government intervenes little in economic decision making? This question is salient for an understanding of *how* posts abroad affect exports: do they serve primarily to help would-be exporters connect with private sector purchasers abroad (i.e., a "networking" role), or do they serve also or even mainly to reduce red-tape associated with government intervention in economic decisions abroad?

To investigate this question, we construct a variable for representation abroad weighted by an index of economic freedom. We find that, using the same data base as Rose, the value of an additional post in an economically less free economy (e.g., China) is worth more in terms of a percentage increase in exports than one in a freer economy (e.g., New Zealand). Encouragingly, this specification of the foreign post variable improves the overall goodness of fit of the conventional gravity equation and increases the significance level of the parameter for posts. We conclude that, all else being equal (which in this instance is a particularly important stipulation, especially as regards the size of the foreign economy), the value of a post abroad is to some extent based on the role of the foreign service in smoothing the interface with foreign governments, over and above any role the post might have in developing private-private networking.

The balance of this Chapter is organized as follows. The next section sets out the basis on which trade promotion programs implemented by posts abroad are understood to promote exports to motivate the analytical distinction between trade facilitation in the private-to-government dimension versus in the private-to-private dimension. The third section describes briefly the construction of the variable to

test this distinction. The fourth section describes the results of the experiment. The fifth section provides some concluding comments.

Situating Trade Promotion in an Economic Framework

International business development activities can be situated in a standard trade theoretical framework. Analysis shows that there exists a strong "home bias" in commerce: commercial transactions across borders are much less likely than within national borders.[3] Home bias is explained by the presence of frictional costs of conducting international commerce, including transport costs, tariffs, non-tariff barriers and others. The business of international business development is essentially the reduction or removal of various types of frictional costs.

The kinds of activities that first come to mind when thinking about facilitation of international commerce address costs embedded in the institutional setting that can be and are addressed by the private sector:

- export financing and risk management, which are addressed by banks;
- contractual aspects of international commercial transactions, which are addressed by legal firms;
- information brokerage, which is provided by various types of consultancies; and
- the deepening of business networks through, for example, sectoral conferences/trade fairs, which are organized by industry associations and other private agents.

Government involvement in these areas thus requires that a *prima facie* case be made that some degree of market failure exists. The *prima facie* case is based on the following:

- The observed "home bias" in commercial transactions is larger than can reasonably be explained by standard economic explanations.

---

[3] The seminal paper in this regard was by John McCallum, "National Borders Matter: Canada-US Regional Trade Patterns," *American Economic Review* No. 3 (1995): 615-623.

- The commercial supply of international business facilitation is likely to be restricted by the same costs that deter international trade in goods and services. In a context of high entry costs, industry structure tends towards fewer large suppliers and higher prices than would be observed in a competitive market. In other words, there is an expectation of an undersupply of private sector facilitation for international business.

- There are positive externalities to an economy from its firms becoming exporters. For example, market knowledge spillovers can bring other home country firms into a market pioneered by one firm. The pioneer firm does not capture these benefits but the home economy more broadly does.

- The archetypal multinational enterprise today is a medium-sized firm with plants and technology partners in several countries selling highly specialized products, with sales in the tens of millions but a large share of the world market.[4] For SMEs, entry costs to the global market are high relative to their resources, which implies that high international market entry costs would be a deterrent at the margin.

On the basis of these arguments, the extensive engagement of government in facilitating international commerce can thus be readily understood. The tools available to government in this set of activities include principally:

- export financing agencies, which fill gaps in export finance, risk management tools, and market intelligence concerning country risk and sectoral trends;

- trade promotion officers, who administer domestic programs to help domestic companies get export ready, and when posted abroad provide local market intelligence and facilitate the development of business-to-business and business-to-government contacts for trade, investment and innovation.

Some costs to international commerce, such as those imposed by government policies, can be addressed through negotiated agree-

---

[4] See Peter F. Drucker, "Trading Places", *The National Interest*, Spring 2005; available on-line at http://www.nationalinterest.org.

ments, either multilaterally through the World Trade Organization or through bilateral or regional free trade agreements (FTAs), and a number of similar instruments such as: bilateral air service agreements (BASAs) which expand trade in air transport services and can thus facilitate international commerce; foreign investment protection agreements (FIPAs) which establish minimum levels of protection for foreign investors against expropriation etc.; and bilateral taxation agreements which remove tax impediments to bilateral investment flows (e.g., double taxation).

Some costs to international commerce are embedded in the institutional and infrastructural context in which commerce takes place. The tools available to the government in these areas of facilitation of international commerce tend to apply primarily to their own domestic economies as opposed to the environment abroad. Actions in this area include measures to improve efficiency of airports, ports and borders in general by reducing red tape in customs procedures, the procedures governing movement of persons and regulatory obstacles to trade. Although it is a bit of stretch, posts could even play a role in some of these areas; for example, in highlighting to domestic agencies unnecessary costs imposed on their host country's business travellers or students applying to their home country's schools. Subsequent action to reduce such costs can directly or indirectly facilitate two-way trade.

And finally there is the concept of the overarching political "relationship" between the two countries. There are various ways in which a well-maintained diplomatic relationship can reduce costs of doing international business.

(a) Building political capital: Developing a friendly political relationship (e.g., support for partners' agendas in areas important to them) can facilitate initiatives to deepen bilateral economic relationships through agreements such as FTAs, etc.

(b) Advocacy: more general diplomatic efforts and advocacy in support of trade deepening to deepen understanding of the benefits of trade can work in the same direction.

(c) Reducing political risk: political risk is a deterrent to companies contemplating international business, whether in committing the capital to develop export markets and/or in investing in a foreign country. Particularly in countries which are less "economically free" and where governments influence business decisions, culti-

vating a good working political relationship can in principle reduce political risk for private firms.

(d) Resolving bilateral irritants: Although trade disputes are launched by private actors, a good political relationship can facilitate the working out of issues underlying trade disputes, including assuring timely completion of regulatory agency procedures necessary to resolve disputes.

The linkages between diplomatic presence and export performance are thus potentially wide-ranging and often indirect. And it is possible that the less visible indirect roles may be even more powerful than the obvious business client-oriented services for exporters if the latter are in fact displacing private sector support rather than complementing it.

One way to shed light on the relative importance of the alternative channels through which posts influence export performance is to consider the degree of economic freedom in the partner country. Arguably, the less economically "free" a country is, the greater the importance assumed by the political relationship and the government-to-government elements of trade facilitation. The trade benefits of these latter elements would be an "add-on" to the benefits from direct trade promotion that focuses on market intelligence, building networks etc.

Adjusting the diplomatic presence variable for economic freedom

We propose a simple adjustment to the variable describing the number of posts that a country has in its trading partner. We weight the posts index developed by Rose with the inverse of the economic freedom index (EFI) published by the Fraser Institute it is annual *Economic Freedom of the World Report*. The Fraser Institute defines economic freedom as "the extent to which one can pursue economic activity without interference from government. Economic freedom is built upon personal choice, voluntary exchange, the right to keep what you earn, and the security of your property rights."[5]

---

[5] Fraser Institute website
http://www.fraserinstitute.ca/economicfreedom/index.asp?snav=ef; accessed June 21, 2005.

Following Rose (2005:6) we estimate the following benchmark equation:.

$$\ln F_{ij} = const. + FM_{ij} \, '(EFI_{j} \, ' \, Average(EFI_{j})) + \ln D_{ij} + \ln GDPC_{i} + \ln GDPC_{j}$$
$$+ \ln POP_{i} + \ln POP_{j} + RTA_{ij} + CurrUni_{ij} + \ln Area_{ij} + ComLang_{ij} + Border_{ij}$$
$$+ Landlocked_{i} + Island_{i} + Colony_{ij} + \varepsilon_{ij}$$

(eq.1)

where

F_{ij} : is the amount of export from i to j.

FM_{ij} : Number of foreign missions of country i in country j.

EFI: Economic Freedom Index.

D_{ij} : distant from i to j.

GDPC: GDP per capita

POP: population

RTA_{ij} : Number of Regional Trade Agreement between i and j

CurrUni: Currency Union (When both countries have same currency, then unit, else zero.)

$Area_{ij}$: Power of areas of export country i and import one j

ComLang: Common Language (English) (When both countries use English, then unit, else zero.)

Border: Border sharing (When both countries share border, then unit, else zero.)

Landlocked: If export country has no sea shore, then unit, else zero.

Island: If export country has no land border, then unit, else zero.

Colony: If both country has colonial links with each other, then unit, else zero.

By construction, the EFI variable will imply a greater contribution to exports from a post in a less economically free country. The real test of the theory thus is whether the weighted posts index has a

higher level of significance compared to the unweighted index and explains more of the variation in export performance.

The Results

We start with Rose's data, obtained from his website. Because the coverage of the Economic Freedom Index is limited to 123 countries, the number of observations is reduced.

First, we estimate a new benchmark equation, using Rose's data for the trimmed down data set covering 123 countries and using his foreign posts index without taking into account the degree of economic freedom.

Table 1
Benchmark results

Variables	Coefficient	t-value
Number of foreign missions	0.08	5.70
Log Distance	-0.56	-12.46
Log Exporter GDP per capita	0.88	31.04
Log Importer GDP per capita	0.92	52.82
Log Exporter Population	1.00	34.57
Log Importer Population	0.99	49.30
RTA	0.91	9.60
Currency Union	-0.58	-3.38
Log Product Area	-0.20	-16.64
Common Language	0.64	8.89
Land Border	1.02	6.40
Landlocked	-0.52	-11.11
Islands	-0.18	-5.01
Summary statistics		
Constant	-35.70	
Number of observations	2622	
R2	0.80	
Root MSE	1.259	

Second, we re-estimate the equation using the foreign posts index weighted by degree of economic freedom. These results are shown in Table 2. As can be seen, the explanatory power of the number of for-

eign missions, importer GDP and common language parameters is slightly increased; however the explanatory power of the remaining variables declines a little. The overall goodness of fit of the model is unchanged.

Table 2
Results with Foreign Posts Index weighted by Economic Freedom

Variables	Coefficient	t-value
Number of foreign missions modified with EFI	0.09	5.76
Log Distance	-0.56	-12.39
Log Exporter GDP per capita	0.88	30.86
Log Importer GDP per capita	0.93	53.96
Log Exporter Population	1.00	34.43
Log Importer Population	0.99	48.39
RTA	0.90	9.58
Currency Union	-0.57	-3.34
Log Product Area	-0.20	-16.64
Common Language	0.64	8.94
Land Border	1.00	6.29
Landlocked	-0.52	-11.08
Islands	-0.18	-4.98
Summary statistics		
Constant	-35.62	
Number of observations	2622	
R2	0.80	
Root MSE	1.260	

We noted some problems in Rose's data set as regards distance and common language. CEPII has alternative data. Distance is defined by the great circle distance between main cities, mainly meaning biggest population, of both countries; common language means the use of the same language, not only English, is used by both nations by at least 9% of each population. The estimation results using these alternative data are shown in Table 3.

Table 3
Re-estimation with CEPII data

Variables	Coefficient	t-value
Number of foreign missions	0.06	4.74
Log Distance (CEPII)	**-0.90**	**-22.49**
Log Exporter GDP per capita	0.87	32.69
Log Importer GDP per capita	0.92	56.25
Log Exporter Population	1.01	37.05
Log Importer Population	0.98	51.60
RTA	0.38	4.13
Currency Union	-0.60	-3.78
Log Product Area	-0.17	-14.50
Common Language (CEPII)	**0.72**	**10.70**
Land Border	0.36	2.37
Landlocked	-0.52	-11.82
Islands	-0.09	-2.74
Summary statistics		
Constant	-32.89	
Number of observations	2622	
R2	0.82	
Root MSE	1.181	

Compared to Table 1, significance level of each coefficient is higher except the coefficients for the number of foreign missions, RTA, product of area, border and islands. Both the impact and the significance level of distance and common language is increased and the overall goodness of fit of the model is also improved somewhat.

Finally, again, we re-estimate the equation using the foreign posts index weighted by degree of economic freedom. These results are shown in Table 4.

As can be seen, the explanatory power of some parameters, i.e., importer GDP per capita, RTA and common language increase. However, the explanatory power of the others declines marginally. As a result, the overall goodness of fit is unchanged.

Table 4
Re-estimation with CEPII data with Foreign Posts Index
weighted by Economic Freedom

Variables	Coefficient	t-value
Number of Foreign Missions modified with EFI	0.07	4.49
Log Distance (CEPII)	**-0.89**	**-22.37**
Log Exporter GDP p/c	0.87	32.55
Log Importer GDP p/c	0.92	57.46
Log Exporter Population	1.01	36.92
Log Importer Population	0.98	50.82
RTA	0.38	4.15
Currency Union	-0.59	-3.73
Log Product Area	-0.17	-14.5
Common Language (CEPII)	**0.72**	**10.76**
Land Border	0.36	2.35
Landlocked	-0.52	-11.78
Islands	-0.09	-2.72
Summary Statistics		
Constant	-32.89	
Number of observations	2622	
R2	0.82	
Root MSE	1.181	

Conclusions

The degree of economic freedom in the partner country plays a statistically significant, if overall modest, role in determining the extent to which posts abroad contribute to a country's export performance. The less economically free a country is, the more important the role of posts.

While it is tempting to draw conclusions from these results about the relative value of an additional consulate in, say, China versus the United States, we would caution against drawing such conclusions on the basis of the present analysis for several reasons:
(a) First, the posts index does not reflect the number of personnel, including commercial officers, in posts abroad.

(b) Second, an additional post in a larger importer will be found to generate a larger volume of trade because the overall trading relationship is larger.

(c) Third, the likelihood of diminishing returns to additional posts is not taken into account. In point of fact, the predicted level of exports into a market in which a country has an unusually large number of posts (e.g., Canada's 20 posts in the United States [6] or Japan's similarly large number of posts, 18 in all, in the United States[7]) is much larger than actual exports.

Accordingly, more in-depth analysis of this issue is required to support such decisions. However, the present analysis does provide support for the proposition that the role of posts is more important in host countries which are less free economically. Thus, when it comes to two broadly similar markets in which a country has a similar number of posts, it is reasonable to conclude that an additional post in the economically less free market would give greater value in promoting exports.

---

[6] Department of Foreign Affairs and International Trade (DFAIT). (2005). *Canadian Government Offices in the U.S.* Retrieved June 23, 2005, from http://www.dfait-maeci.gc.ca/can-am/menu-en.asp?mid=1&cat=16

[7] Ministry of Foreign Affairs (MOFA), Japan. (2005). Zaigai Koukan List. (in Japanese). Retrieved June 23, 2005, from http://www.mofa.go.jp/mofaj/annai/zaigai/list/n_ame/usa.html

Multilateralism, the Club Scene or Entrepreneurial Activism? Choices in an Age of Change

John M. Curtis[*]

The contemporary discussion of the international institutional system engages, explicitly or implicitly, a range of extraordinarily broad themes. First, there are varying levels of dissatisfaction with the multilateral economic institutions conceived at Bretton Woods against a background of deep concern over the viability of the mother ship itself, the United Nations. Second, there is a search for new, intermediate institutional structures such as the "G-somethings" (or, at the national leaders' level, the "L-something") to act as catalysts or brokers on issues stalled at the multilateral level. Third, there is a concern over the ways and means for continued bilateral collaboration between friendly nations based on shared values, interests and obligations in global affairs.

This discussion is taking place in a context that brings together the political with the economic and the security dimensions of international relations in today's world. I should therefore ask for the reader's indulgence: the remarks below are somewhat speculative and necessarily in a personal rather than governmental capacity. My main messages are as follows:

First, Canada, one of the most open of the G-7 countries, has an uncommonly deep interest in maintaining a well-functioning global trading system and stable currency frame-

works; by the same token, Canada, while playing only a limited role in the main imbalances and tensions in the international trade and finance systems, has a deep interest in being part of the solution intellectually and, to the extent possible, operationally.

Second, the traditional exclusive "club scene", flattering as it is to be a member, might not hold out the opportunities for medium-sized powers like Canada that it did in the past. For Canada, this reflects largely geo-economic inevitability. At the same time, there might be more life and promise in the club scene than many might believe, but primarily in the context of the multilateral institutions: I refer in particular to the role of new informal groupings that organize to address the issues of the day in context of institutions such as the World Trade Organization. These provide, I would argue, ample opportunity for middle-power states, through entrepreneurial activism, to make a difference.

Third, if John Maynard Keynes was correct in his restatement of the aphorism that the "pen is mightier than the sword" when he famously said that those who fancy themselves men of action are in reality the slaves of a defunct economist, the first step in identifying solutions to the current global economic problems might not lie in determining the invitation list to the next G-something, but in identifying the defunct economist in whose thrall we currently toil—and I would add, the defunct scribbler on international relations whose influence in that field appears to be no less great!

The primacy of ideas over institutions

Let me start with the last theme first. In economic terms, the global community, frankly, is in uncharted waters.

The world has become trade dependent to a degree that I believe it safe to say is, notwithstanding the high degree of openness attained by 1913, unprecedented. Counting both cross-border trade in goods and services and the sales of foreign affiliates—Mode 3 or "commercial presence" trade in WTO jargon—international commerce is roughly equivalent to 2/3 of

global GDP. Given the extensive degree of fragmentation of production across borders—or, put another way, the integration of economies through outsourcing and trade in intermediate goods (the so-called "global value chain")—our economies are if anything even more interdependent than that figure suggests. If one were to trace the pedigree of even the most humble of domestic products intended for domestic consumption, it would be hard to find one that does not have some degree of international content, whether in terms of components, intellectual property, or the capital equipment or financing used in its production. When a Canadian barber gives a Canadian customer a haircut, a quintessentially "domestic" economic transaction, chances are that either the scissors, the comb, the electric razor used to trim the sideburns or perhaps even the barbershop chair are foreign products, most likely from China!

Yet the price structure that internationally links this web of production and consumption on which we so depend works through a system of exchange rates which since the breakdown of the Bretton Woods system in 1971-1973 has exhibited a degree of volatility that is hard to understand in terms of any current theory of exchange rate determination. It is not unusual these days for the Canada-US exchange rate to move by a full cent in a trading day—annualized, this represents a rate of change (calculated over some 250 trading days) of well over 2,000 per cent. While much of this is high frequency "noise" in the price system with a powerful tendency for excessive movements in one direction one day to be corrected by offsetting movements the next, the extent of movement over periods of months and years has confounded experts and often defied explanation, even with 20-20 hindsight. Common sense suggests that, as firms and countries adapt to cope with volatility, through various forms of hedging for example, the latitude for volatility increases. And periodically, such as during the Asian Crisis, the volatility breaks completely out of control and acquires an elemental destructiveness that, in the economic sphere, resembles a force 5 hurricane breaking on land.

Further, the system—which in theory is supposed to prevent the build-up of imbalances, being self-correcting in na-

ture—is generating imbalances that are not only puzzling to economists but increasingly disturbing to the markets themselves. Of course, the most significant imbalance currently is the US current account deficit, which has been 4 percent or larger since 2000, has recently surpassed six percent of GDP, and is still growing. Conventional forecasts project it to remain in this range for the foreseeable future. The US net international asset position has deteriorated from a positive net balance for most of the postwar period to a negative balance of about US$2.5 trillion at the end of 2004. While the world works on an international dollar standard, which means that the US faces no hard constraint since it borrows in terms of its own currency and can always pay its debt by printing more dollars, the US still does face an uncertain constraint in terms of the willingness of foreigners to hold its fiat money. Similarly, while a US dollar depreciation could correct its external asset imbalance, the foreigners holding net positions in US-denominated assets face a substantial capital loss if that happens. While to a large extent the resolution to this issue lies in a correction of the US internal savings-investment balance, insofar as part of the solution is to shift the burden internationally, the system faces an interesting adjustment, to say the least.

Thirdly, problems in the development process have emerged that our conventional economic theories simply did not anticipate. The basic theory of international trade, that of comparative advantage, does not suggest that some nations will be left behind; quite the reverse—even if a given state has an absolute competitive disadvantage in the production of every conceivable product, it would have a comparative advantage in some product and so would be able to participate in the global trading system. In a similar vein, going by the basic concepts of scarcity and diminishing returns, capital-scarce regions of the globe ought to offer high returns to capital while capital-rich countries ought to offer comparatively lower rates of return. But the direction of flow of international capital is not in line with these theoretical expectations. To be sure, theoretical refinements have been introduced to explain actual patterns and "save" the basic theory; the problem is that the "wrinkles" to the

214

theories have to do enough heavy lifting that they stand the expectation of the basic theory on it head! That this is too much to ask of basic theory should have long been obvious to all.

We may choose to express dissatisfaction with the multilateral institutions responsible for the global systems of trade, finance and promoting development; but to me that makes about as much sense as shooting the messenger for bad news or firing the manager of a sports team because the players are truly awful. Put another way, the IMF and World Bank did not invent the so-called Washington Consensus—this was distilled from the experience of successful emerging markets and applied as a prescription for development in other countries. The problem was that, like other prescriptions previously used, it did not seem to consistently work as a prescription. Similarly, the WTO is merely the forum in which countries negotiate the rules of the road or rules of the game for important aspects of international commerce and their own terms of engagement; changing the forum will not in any way change the nature of the issues faced by its members, which today include most of the international community.

To be sure, sometimes an entrenched bureaucracy is part of the problem in preventing a change from failed solutions to more promising approaches; and in terms of the metaphor used earlier, sometimes the manager of a sports team is part of the problem. However, it is not obvious to me that presently the problems lie so much in the institutions as in the state of our understanding of the dynamics of the systems of international commerce and of the process of economic development.

Surprising life in the multilateral system

Two recent events have helped rather dramatically to confound trend watchers perceiving an unstoppable momentum towards regionalism, bilateralism and even unilateralism. I refer to the Duelfer Report on the success of the UN in its arms inspections in Iraq and the July 2004 agreement in the WTO on a negotiating framework for the Doha Round.

There are certainly costs to going the multilateral route. Easy things are harder and more complicated to achieve; however, the really big, hard things are often only attainable through that route. For example, in the world of international trade, a process such as a "Single Undertaking" in the context of a broad round that puts enough diverse things on the table to form the basis of positive-sum or "win-win" outcomes for all participants might well be the only way to break through on otherwise intractable problems such as agricultural subsidies.

Further, whatever the frustrations, the multilateral process has the serious advantage of legitimacy that smaller groups simply do not. The legitimacy conferred by multilateral agreement in the security sphere is clear. But this is also true in the world of trade. For example, the WTO, through Article XXIV in the GATT and Article V in the GATS, establishes the conditions under which regional trade agreements are consistent with multilateral obligations. Thus, not only does a multilateral system help minimize the downsides of regional trade arrangements—such as trade diversion –but it also defuses the potential for such regional agreements to spiral out of control into economic conflict between rival blocs.

Generally, middle powers need little counsel or persuasion regarding the benefits of a sound multilateral system. However, in Canada's case, it has been spoiled by long having had an insider's position through membership in the G-7. But seen in historical perspective, clubs such as the G-7 form to deal with particular issues and cannot be taken complacently as being permanent fixtures.

The origins of the G-7 date back to the breakdown of the Bretton Woods system of fixed exchange rates. On March 25, 1973, US Treasury Secretary George Shultz invited the finance ministers of France, Great Britain and Germany to an informal discussion on the international monetary issues caused by the US's decision to abandon the gold standard a year and a half earlier. This initial meeting led to a series of informal meetings over the following months of the four plus the Japanese finance minister. Thus emerged the original "Group of Five" or G-5. Two of the finance ministers at these meetings were Valéry

Giscard d'Estaing and Helmut Schmidt who subsequently became President and Chancellor of France and Germany respectively. It was the initiative of Giscard to replay the experience of these informal finance ministers' meetings at his new Presidential level; thus was conceived the first summit of the world's leading industrialized nations, at the Château de Rambouillet, on November 17, 1975. Giscard of course invited Italy and this in turn led to Canada being extended an invitation the following year.

The salient point here is that the make up of the group was dictated by the key issue: the G-5 brought together the main parties that had been central in the breakdown of Bretton Woods. As will be recalled, it was the decision of France and Great Britain to request, in accordance with the terms of Bretton Woods, an exchange of gold for excess dollar holding which prompted the Nixon Measures of August 1971 that abrogated that agreement. Germany and Japan meanwhile represented the two countries with the most important roles to play in the international balance of payments position of the US in terms of their weight in both the trading and currency systems.

It is instructive in light of that experience to consider how the more recent July 2004 agreement in the Doha Round negotiations came about. The emerging WTO dynamic included first the formation of two blocking coalitions: of key emerging markets (the new G-20), and of the poor countries (the G-90). These coalitions had earlier balked at the offer from the industrialized countries that was on the table at the WTO Ministerial meeting in Cancun in September 2003. The collapse of the Cancun Ministerial then prompted a flurry of activity including the emergence of a new "non-group" of five countries, the so-called "Five Interested Parties" which "worked the room" in Geneva and succeeded both in terms of achieving agreement on some issues and constructively papering over divisions in others areas to allow the Doha Round to proceed. The five included the US, the EU, seen as "representing" the industrialized countries, Australia "representing" the Cairns Group of agricultural exporters and India and Brazil which were the leading activists in the G-20. The invited group thus straddled the key divide that had

scuttled the WTO Ministerial conference at Cancun. A parallel to the original Finance G-5 is clearly suggested in this process.

For institutional architects, the fact that the July 2004 WTO agreement was not brokered by the semi-permanent clubs was perhaps a setback of sorts; for enthusiasts of self-organizing "emergence", a feature of complex adaptive systems, in this case in clear support of multilateralism, this was affirmation of the importance of informality and of the need to be prepared for surprising developments which such systems routinely conjure up.

Implications for middle powers

The club scene in international relations is not permanent; just as with night clubs, some become less "fashionable"—or in our context less relevant—while others emerge. Thus, even as the WTO scene has been energized by the formation of the G-20 of dynamic emerging markets, the "Five Interested Parties", and the G-90, "traditional" alliances are gradually disappearing from the scene, including the Quad and the Cairns Group. Recent years have also witnessed a flurry of "mini-Ministerials" that involved a changing mix of WTO Members meeting for various WTO-related reasons, but generally serving to inform and persuade various parties to the negotiations. In the world of finance, we see a similar evolution of institutional structures with the invitation to China to join on occasion the G-7 Finance Ministers discussions on international financial matters. Further, a proposal was recently mooted by the US Council on Foreign Relations for a new Quad of the US, the EU, China and Japan to serve as the core group to manage the "big picture" in the international finance domain.

What do all these developments imply for middle powers such as Canada which might not always be part of the problem that drives the structure of emerging clubs? One answer is that the key to being part of the club scene is to be energetically involved—including by providing new perspectives and innovative ideas—in the multilateral process which is the breeding ground for the clubs. Perhaps a second way is to build on Den-

mark's "Copenhagen Consensus" process, an issue-driven invitational approach to crack large, difficult issues.

The world has moved a long way since George Schultz launched the process that brought Canada into the G-7. It is timely for us now to look forward.

The Convergence of
Law and Economics:
Implications for the World of Trade

John M. Curtis [*]

Introduction

Law and economics are two distinct disciplines, rooted in very different soils, developed to solve different problems encountered by evolving societies. Each has created its own characteristic traditions. Each draws on its own set of epistemological tools. Law is largely deductive, reasoning from principle and precedent to establish new precedents (and perhaps even new principles); reality is tested against the law and if necessary is constrained to conform to the law. Economics is inductive, discerning systematic elements in the chaotic froth of economic life that are then labeled "economic laws"; these laws are tested against reality and, when found wanting, it is the law that is reshaped or restated to conform with the reality (although sometimes ideologues in the field of economics do precisely the reverse!)

Yet law and economics also have fundamental similarities. Both are concerned, for example, with procedural and substantive "fairness". In law, we have the well-known concepts of *due process* and *mete justice*. It is less well understood that there are essentially similar concepts at play in economics. For example, the notion of a market transaction incorporates the notion of procedural fairness in the sense that the transaction must be between a willing seller and a willing buyer with no coercion ap-

[*] John M. Curtis is Chief Economist, Foreign Affairs and International Trade Canada. This Chapter was originally presented as a guest lecture at the University of British Columbia. The views expressed are the views of the author and are not to be attributed to any institutions with which he is affiliated.

plied in the setting of the terms and conditions of the transaction. The notion of a market transaction also incorporates the notion of substantive fairness in that such a transaction is deemed to represent an exchange of value for value—hence the notion of "fair market value". Indeed, the very possibility of exchange is seen as allowing each party to be better off than before the trade.

In terms of their roles in societal development, law and economics are in many ways symbionts. Economics, if it be given its due, helps to drive the evolution of our economies. But ever larger, more complex, socioeconomic arrangements generate new problems to be overcome in order that their benefits can be accessed and indeed sustained. Law helps to establish the institutional frameworks that regulate the new forms of social interaction stimulated by developments in the economic sphere. Law provides the disciplines that mediate the frictions that inevitably emerge. And, by injecting its own norms and values into the socioeconomic DNA, law in some sense helps shape the emergent socioeconomic form, allowing it to serve as a more viable base for the erection of the next stages of the socioeconomic edifice.

Let me give a few practical examples from various fields of economics and law.

The term *legal tender* preserves the memory of the transition from metallic to fiat money. As economic growth generated increasingly large incomes and transactions volumes, these outstripped the capacity of the mines and mints to supply what once was known as the "coin of the realm" to meet the resulting demand for a medium of exchange. The institutional development of fiat money was made possible by the legal requirement that it be accepted in fulfillment of debts. A gold sovereign had its own intrinsic value; no law was needed to require its acceptance. Not so with paper. Trust amongst a small group known to each other, as in the Italian financial brokers whose practice of borrowing and lending money through IOUs issued on benches in city squares gave rise to the term "bank" (from the Italian for "bench"), also has its limits. To function in the anonymity of an evolved marketplace, fiat money proved necessary. But eco-

nomics did not rest there; it sought out further innovations. Thus was called into existence the vast superstructure of financial institutions and markets that now forms an essential part of the livelihood of the great financial capitals. And of course, in turn, these economic developments spawned another layer of legal development: thus do we arrive at the full complexity of modern banking and securities laws.

The enforcement of contracts, a quintessentially legal concept, is seen by the Peruvian economist Hernan De Soto as an essential element of economic development. He observes that gold is traded on the London Exchange and on the streets of Lima. What distinguishes London, he argues, is the large volumes and low cost of trading, which are made possible by the legal framework in which the London Gold Exchange operates which gives confidence to traders. One market flourishes and the other does not because of the legal infrastructure that the one possesses and the other does not. And here as well, economics continues to build, once a sound basis is established: witness the activities of commodities exchanges which produce exotic contracts such as puts and calls in pork belly futures contracts and so forth.

The notion of private property is couched and circumscribed in legal language. Yet it is at the heart of economic frameworks. China's modern economic miracle is widely held to have been kicked off by a policy change that effectively created property rights in the agricultural sphere—the so-called family responsibility system. In other words, in a socioeconomic system without property rights, the introduction of such rights unleashes economic energies previously held in check as we have seen since the reforms that began in 1978.

Conversely, economics concerns itself deeply with normative, indeed moral, issues and thus invades territory that the law might consider properly within its own domain. In the early development of economics, there was no border between it and philosophy. David Hume and Adam Smith had one foot firmly planted in each discipline; their economics was situated in a moral universe. John Stuart Mill's reworking of classical economics was in good measure based on the notion that the distri-

bution of wealth was not based on natural law but on the "laws and customs" of a society—in other words on its institutions. Friedrich von Hayek's later defence of the market system was, as we were recently reminded by Amartya Sen in a Financial Times op-ed[1], based on the individual freedoms which the market system delivered, not on the wealth that it might generate. Here we have economists setting out the economic basis of individual liberty and fair distribution of a nation's wealth. The contribution of John Maynard Keynes in articulating the role of government in regulating economic activity as a guarantor of an implicit social contract, at the expense of its own treasure rather than for the increase thereof, needs little description.

Law and economics are not, of course, the only disciplines at play in shaping the evolution of modern socioeconomic frameworks. At a minimum, one must add political science to the mix. But the interface between law and economics is of particular interest in this Chapter. And it is this interface that I should like to explore more deeply with reference to the world of trade.

The Convergence of Law and Economics in Trade

International trade can legitimately be described as being at the very root of modern economic thought from its first stirrings in the European Renaissance to its first full flowering in the Enlightenment. For example, one of the very first recognizably modern economic texts was written by Thomas Mun of the East India Company in defence of the Company's exporting capital in order to transact business abroad. The title of his work, *England's Treasure by Foreign Trade*, written ca. 1630, speaks for itself. It provides an essentially complete description of the principles behind the international balance of payments, which would require only up-dating of the language to sound perfectly modern. To be sure, much of the early pre-Adam Smith theorizing was in defence of mercantilist practices; but equally, it

---

224

must be pointed out, the major seminal works of modern economics, including that of Smith himself, were very much motivated as an attack on mercantilism. David Ricardo's articulation of the theory of rent, for which he is primarily known today, found its first application in the debate over Britain's Corn Laws. Ricardo's arguments helped make the case for the abolition of these agricultural protections, which in turn had (arguably) much to do with the Britain's emergence as the pre-eminent industrial power of the 19th Century.

And international trade was also at the very root of the establishment of the post-World War II economic order at Bretton Woods: it was the damage wrought by protectionism in the 1930s that motivated the framers of the agreements that established the International Monetary Fund, the World Bank and the provisional General Agreement on Tariffs and Trade (GATT) —the latter as stand-in for the proposed but never approved International Trade Organization.[2]

The early postwar economic order was not much concerned with institutions, however. *Laissez faire,* as qualified by Keynes and the Keynesians, relied on market forces with a judicious admixture of government macroeconomic intervention to regulate economic growth and propel development. It was only when this approach stumbled in the 1970s with "stagflation" (a combination of economic stagnation and high inflation that was not anticipated under the Keynesian economic model) at home, major currency and commodity price fluctuations and trade imbalances internationally, not to mention an unexpected cessation and even reversal of broad-based development in what came to be called the Third World, that a fundamental rethinking of economics was prompted.

And in good measure, it was to the institutional frameworks in which economies functioned to which economists turned their attention. One of these institutional features was money; the result was the monetarist revolution in macroeconomics. In development economics, successive waves of reform were

---

[2] Diebold, W. J. (1952). "The End of the I.T.O.", *Essays in International Finance No.16*. Princeton: Princeton University. October: 1-37.

launched by the institutional architects sitting in the international financial institutions--infrastructure, education, family planning, rule of law and legal/technical institutions such as bankruptcy laws and latterly political institutions themselves each of which came to be seen in their turn as the "silver bullet" to restore development. In the world of trade, it is perhaps no surprise that the Tokyo Round, which was negotiated in that pivotal decade of the 1970s, was noted less for tariff cuts than for several supplementary agreements that reached for the first time inside borders to address institutional matters. These supplementary agreements included the codes on dumping, government procurement, standards and subsidies, as well as strengthened procedures that substantially expanded GATT's role in resolving trade disputes.

Trade lawyers soon had plenty of grist for their mills. In 1979, the same time that saw the conclusion of the Tokyo Round, the United States transferred the administration of important elements of its trade remedy laws from the Treasury Department, which tended not to see injury in alleged dumping cases, to the Commerce Department, which appeared to have had no trouble finding dumping or injury—even in much-celebrated cases where exporters selling into the United States charged higher prices in that market than at home! The "legalization" of international trade was thus well underway well before the Uruguay Round which many observers tend to identify as the watershed event in that regard.

Trade scholars tend to see a sort of natural, inevitable pattern in the evolution of trade policy from a primarily economic discipline, focused on addressing the distortions caused by tariffs and other trade barriers to prices and consequently to the basis for competition between domestic and imported goods, towards a legalistic discipline epitomized by the dispute settlement panel operating in an unmistakably legal setting, replete with judicial procedures and practices. In the WTO agreements, many trade scholars see a nascent constitution for the world economy; in the WTO's Appellate Body, they see a semblance of a Supreme Court. Even the controversy that has embroiled the WTO—that of "legitimacy"—has been described by trade

lawyer and theorist Joseph Weiler, as "part of the standard vocabulary of court watching".[3]

Trade scholars tend to attribute this pattern of evolution as being due to deepening "integration": as trade barriers fall, sharper price competition causes states to fall back on non-tariff measures to protect domestic commercial interests. Trade rules then necessarily reach "inside the border" to root out these measures. In so doing, these trade rules can start to affect even measures that were not erected for protectionist purposes—which are conceded, perhaps somewhat grudgingly in the case of the trade policy community, to be "legitimate": Sylvia Ostry's "system friction". However, many economists' paradigm allows them to interpret differences in domestic laws and practices as not being due to democratic choice but as mere departures from the most efficient way to organize society. If trade rules reach in and force tax regimes into line with the most efficient international regime, they would argue, so much the better.

In making such interpretations, trade scholars and economists more generally are simply echoing the well-rehearsed arguments established by the free traders critiquing the mercantilists. And if they are right, we are seeing not so much the convergence of law and economics in trade, but rather the passing of the torch from economists as the advocates of free trade to the lawyers—implicitly, the role that economics can play in establishing a liberal international trading order has largely been played or, put another way, there is little powder that economists have not spent. The remaining task is to ensure that the gains made by application of economic principles are not eroded and that the liberal trading order is not undermined, a task that falls essentially to the legal domain.

But I am not convinced that this interpretation is correct.

First, at the microeconomic level, the organization of society features an interplay between economics (especially in terms of the incentive structures which characterize a socioeco-

---

[3] J.H.H. Weiler, "The Rule of Lawyers and the Ethos of Diplomats: Reflections on the Internal and External Legitimacy of WTO Dispute Settlement", 35(2) Journal of World Trade 2001, 191-207, 193.

nomic construct), law (especially the constraints on behaviour which laws establish), and politics (in recognition that relations among people cannot be reduced to those aspects that are tractable in economic models!). If we wish to understand the impact of trade agreements, surely the analytical approaches that we adopt must integrate these three perspectives.

Second, at the macroeconomic level, the 1970s which saw the first major wave of legalization in the trade system also witnessed a remarkable upheaval in the international price system. Waves of instability have subsequently rolled through the global macroeconomic system. These are manifest in shifting external imbalances, wide swings in real exchange rates, the emergence and deflation of bubbles, and regional crises in various parts of the world. From the anti-dumping literature, a fairly persuasive case can be built that much of the activity in this domain is prompted by macroeconomic developments, such as real exchange rate swings. In other words, the legalization of the trade system was driven by macroeconomic dislocations: arguably, responses in the legal domain require a complementary and nuanced economic analytical input.

Conclusion

I see accordingly much scope in the coming years to fruitfully integrate law and economics, together with political science, both at:

(a) the micro level, in understanding how, in shorthand, incentives, constraints, and customs combine to shape the evolution of our societies and economies; and

(b) the macro level, to ensure that the rules-based framework builds in appropriate flexibility to accommodate the disequilibria in which our economies continually find themselves, while at the same time maintaining the disciplines on behaviour that are necessary to promote cooperative behaviour.

In practical terms, that might have several general implications:

One would be the development of interdisciplinary institutes within academia integrating law, economics and political science, where basic pre-policy micro analysis of the sort that I have described can be furthered.

Second would be the professionalization of policy departments within government, with high-level "receptor" capacity in these three areas to absorb and apply the lessons learned in research and analysis in academia—and to stimulate that research by formulating questions and issues on which research would be most valuable.

Third and finally would be interdisciplinary support for policy development and issue management within government—including for example of trade disputes and of trade agreements.

This clearly is the future.

Foreign Affairs and
International Trade Canada

Affaires étrangères et
Commerce international Canada

Canadä

Government
Publications

CA1
EA
T66

Trade Policy
Research
2006

Trade Policy
Research
2006

Dan Ciuriak
Editor

Table of Contents

Foreword

This volume brings together the results of analysis and policy research undertaken within, on behalf of, or in collaboration with Foreign Affairs and International Trade Canada over the past year. Launched in 2001 as part of the response to the Government of Canada's *Policy Research Initiative*, a government-wide effort to re-create and expand its research capacity, the *Trade Policy Research* series is now in its sixth edition.

Previous volumes have traced the debate in trade policy circles since the watershed developments at the 1999 WTO Ministerial in Seattle, following the progress of the Doha Round, touching on topical issues, including the surge in regional trade agreements, and showcasing research and analysis conducted within the Government of Canada on various aspects of trade policy and economic globalization more generally.

This year's volume continues in that vein. Part I provides a report on a roundtable discussion by international trade experts on the prospects for the Doha Round, taking into account the geopolitical and international macroeconomic context, as well as looming developments on the political calendar.

Part II compiles the papers presented at the conference *Integrative Trade between Canada and the United States—Policy Implications*, organized by Carleton University's Centre for Trade Policy and Law, Ottawa, December 6, 2006. The papers in this part examine the implications of the re-shaping of international commerce through the on-going fragmentation of the production process for our conceptual understanding of trade and investment; the issues posed for statistical agencies in grappling with the changing international industrial landscape; and the implications of these developments for the trade policy community.

Part III includes two papers addressing regional trade issues: an assessment of the impact of trade with Canada on US state-level jobs and output, updating and expanding an earlier study on this theme included in *Trade Policy Research 2004*;

and an assessment of the impact of free trade agreements on Canada's automotive sector.

Through this volume, Foreign Affairs and International Trade Canada seeks to continue to contribute actively to the discussion concerning the role of international trade and investment in Canada's economy and in the global economy more generally, to continue to work in the spirit of the broader commitment of the Government of Canada to stimulate the development of its applied research capacity, and to further develop links with professional and academic researchers in the field of international commerce.

This volume was produced under the guidance of Anthony F. Burger, Chief Economist, Foreign Affairs and International Trade Canada.

The Editor
Ottawa
March, 2007

Part I

Multilateral Trade Issues

A New Realism in the Doha Round?
A Roundtable Discussion

John M. Curtis and Dan Ciuriak[*]

On March 2-3, 2006, a group of leading observers of the international trade and investment scene gathered in Ottawa for an informal discussion of the prospects for the Doha Development Agenda in view of the progress in negotiations made at the Sixth Ministerial Conference of the World Trade Organization (WTO) at Hong Kong, China, December 2005. The roundtable discussion was sponsored by the Centre for International Governance Innovation at the University of Waterloo and the International Development Research Centre, in coordination with the Department of Foreign Affairs and International Trade. The talks focused on the progress of negotiations and the prospects for an ambitious outcome, taking into account the geopolitical and international macroeconomic context, as well as looming developments on the political calendar. This note represents the Chair's thematic summary of the discussions. As these were held under Chatham House rules, no attribution is given. Responsibility for the interpretation of the discussion rests entirely with the authors. Although the prospects for a successful conclusion of the Round remain in doubt, the analysis is still relevant.

Introduction

With the benefit of now a half-decade's worth of perspective, it is clear that the Doha Development Agenda was pre-destined to delay and quite possibly to modest results.

Trade agreements have commercial objectives; but the Doha Round was launched in good measure as an international political response to 9/11. At the launch date of what became known as the Doha Development Agenda in November 2001, the commitments of the Uruguay Round had not been fully implemented, let alone absorbed. The technical groundwork for a new Round had not been laid; movement on the built-in agenda mandated in the agreement that concluded the Uruguay Round had been negligible. And the major looming challenges for world trade were to absorb the impact of the then imminent accession to the World Trade Organization (WTO) of China and the expiry of the textiles and clothing agreement.

Contextually, the Doha Round was launched at a point in time when the so-called Washington Consensus on economic policy was unraveling in the wake of a stunning series of emerging market crises that had begun with the Mexican financial crisis in 1994. Governance issues were literally exploding with anti-globalization protests at one international venue after another.

And two of the key issues in the Round were to be development and agriculture; as one indicator of the difficulties posed by these issues, the need to address them in the multilateral trade system was first noted in a report to Members of the General Agreement on Tariffs and Trade (GATT) back in 1958 by a panel of leading experts chaired by Gottfried Haberler. Four rounds have been completed since then; while some disciplines were adopted as part of the Uruguay Round outcome, agriculture and development remain major issues.

Complicating matters was the fact that some developing countries had come to feel that they had been, in the words of one observer at the Roundtable, "ambushed" in the Uruguay Round and were seeking a rebalancing of the results of that Round (which would of course not be in the interests of con-

stituencies in the industrialized world such as the pharmaceuticals that had pushed for the Uruguay Round). Nor did it help that the first concrete steps on agriculture after Doha had been backward—the European Union's extension of the Common Agricultural Policy (CAP) in 2003 and the US Farm Bill of 2002 which expanded US agricultural support quite substantially.

Given the scale of the challenges, it is at least arguable that the Round progressed, all in all, rather well through its first four years—and indeed not out of line with what in retrospect would have been reasonable expectations. Consistent with this expressed optimistic view, the hard slogging in Geneva and various Ministerial meetings (plenary and mini) might well have served to bring expectations into line with a feasible outcome, finally allowing the Round to enter what one observer termed an "Age of Realism" in which the final moves toward agreement might be quickly and decisively taken.

Or not. Contemporary assessment of major events is hampered by the stubborn indeterminacy of the future—which in small part at least reflects the fact that contemporary assessment is itself part of the analytical feedback that plays into ultimate outcomes.

The following is a synthesis of the thoughts and views of close observers of the trade scene as regards the state of the multilateral trade negotiations and more generally of the global trade scene, as they were put forward and discussed in Ottawa shortly following the Sixth Ministerial Conference of the World Trade Organization (WTO) at Hong Kong, China, December 2005.

Is the Round still "doable"?

At the roundtable, it was argued that the outcome of the Doha Round negotiation turns on several inter-related questions:
- whether there is a persuasive commercial case for the round;
- whether a deal can be configured that meets the commercial objectives, with a reasonable balance of negotiating gains and concessions for all Members, and that is at the same time politically feasible; and

- whether such a deal can be put together in time—which for practical purposes means by July 1, 2007 when US Trade Promotion Authority (TPA) expires, or in such additional time as might be provided by an extension of the TPA?

Is there a commercial case for the Round?

Who is interested in the Round? In response to this question, it was noted that the lobbyists do have clients in a number of sectors and that all their clients want a big result from the Round. So there is business interest. This may be more apparent in the United States where the trade policy process is more directly business driven than in Europe, where it is very hard and time consuming for interest groups to press their views, given the layering of institutions (national governments, European Business Associations, and the Brussels bureaucracies).

The breadth and intensity of the interest is not clear, however. Some developing country observers, for example, see an "interest deficit" in the OECD countries—that is, there is no interest in opening up, especially in agriculture. Thus, it was noted, the agreement at Hong Kong to limit sensitive sectors to three percent of all tariff lines provided no assurance of market access gains since developing countries often have exports concentrated in a handful of tariff categories. The three percent carve-out could cover *all* areas of interest to many developing countries, it was suggested.

And some countries have diametrically opposed interests to improved market access—their concern is preference erosion, a major consideration for a large number of the poorest countries.

Accordingly, while it was argued that there is "money in the Round", when one tries to pull together a comprehensive perspective on interest in the Round, the result is a rather confused and not totally persuasive picture.

The political jigsaw puzzle

That being said, it was suggested that the shape of a deal that is do-able in commercial terms is reasonably well understood. In-

deed, one observer argued, if the trade negotiators from the top twenty or so trading economies were to write down their view of such a deal, there would be a high degree of convergence of views. The question is whether we can get there *politically*, domestically and internationally – i.e., is there a feasible solution in political economy terms?

On the one hand, a deal that would bring the then 149 WTO Members into the fold would, it was argued, maximize the pain for the advanced countries, with greater agricultural reform in the EU and the US than had been forthcoming to date. It would also require more in the way of market access concessions from the big emerging markets than has been put on the table so far. On the other hand, a small deal would reduce the pain for the advanced countries but would not be saleable either to the US Congress or to developing countries, particularly the least developed.

Much was seen as depending on the United States and the European Union.

The situation in the United States was seen as problematic if the July 1, 2007 TPA expiry deadline were not met since the case for extension was not there. The gloomy view was that, with the Administration's approval rating in "free fall", the Democrats had no incentive to bail it out. But some held out hope that, by January 2007, with Congress re-shaped by mid-term elections and the unfolding of political and economic events, a case for extension might be made. "Trade votes, it was pointed out, have not been won on economic arguments alone: Foreign policy and security got us into this mess and will get us out of this mess". In support of this view, it was noted that, while the declining polls for the Administration have stimulated protectionist rhetoric in Congress, potential presidential candidates had largely stayed "above the fray" on trade protectionism. As well, recalling that the Uruguay Round agreement was reached shortly after the United States brought in Robert S. Strauss, an individual with a reputation as a "closer", to hammer out a deal, some expressed hopes that Ambassador Portman, an individual

who was popular in Congress, might similarly be the man of the hour. [1]

As regards the European Union, some questioned whether the chief EU negotiator could actually go beyond the still-inadequate package on the table: Is the EU negotiating posture a charade, it was rhetorically asked? If not, what does Commissioner Mandelson have to do to get a change? In response, it was pointed out that, while the EU negotiating process is indeed cumbersome, the history of the Uruguay Round shows that movement is possible. The 1992 reforms to the Common Agricultural Policy (CAP) engineered by Commissioner McSherry made the Uruguay Round agreement possible, as Roundtable participants were reminded. The McSherry plan, it was observed, arose as a result of international pressure placed on Europe during the Uruguay Round (Blair House being the key meeting that set up the EU move). Importantly, it was further argued, the farm situation in France, which has been the key stumbling block for CAP reform, is changing. Popular sympathy for the CAP had historically been based on the notion that it helped small farmers; in reality, the CAP funds flow mainly to large farmers (In fact, it was indicated, the Queen of England and the Prince of Monaco have been identified as major beneficiaries). This is becoming increasingly understood, which is causing a shift in public opinion. The European poor are the main losers from the CAP and consumer attitudes against high food prices are hardening. At the same time, French farmers are becoming aware that they can become competitive (including through greater use of genetically modified crops). The French support freer trade by a margin of 60-40, it was asserted, preferring the movement of goods to the movement of people. The EU debate is thus changing and chances of a breakthrough should not to be entirely written off. However, timing is uncertain: as is the case with tectonic plates, pressure builds up and then there are sudden shifts, the timing of which is hard to predict.

---

[1] Editor's note: These hopes were dashed shortly after the Roundtable when Ambassador Portman was appointed Director of the Office of Management and Budget in June 2006.

The question of time

The time required to negotiate a deal must be understood as being in part a function of the effort provided. Thus, it was argued, with an injection of some new energy and, given sufficient effort, a big deal could be put together in a short time.

However, given where the negotiations were in the immediate aftermath of the Hong Kong Ministerial, some saw the unofficial deadline of TPA expiry as, for practical purposes, having already been missed. If so, it was argued, public perceptions and the negotiating dynamic would depend heavily on whether the US Administration can get TPA extension. With TPA extension in hand, July 2007 would then be seen as the midpoint of an extended but live Round; without extension, talks might continue but July 2007 would be widely interpreted as marking the *de facto* failure of the Round: in the words of one observer, "Without the TPA deadline, things would stop."

To some observers it seemed quite extraordinary that the whole exercise *should* depend on US trade negotiating authority. Amongst the developing countries, it was noted, there is unhappiness with being in effect "blackmailed" by TPA expiry as a deadline for the negotiations.

But others argued that it was important to take advantage of the deadline provided by TPA expiry for the good of the multilateral trade system. Regionalism, it was suggested, is "at the gates".

Rounds do get done

While there was no clarity regarding the sufficiency of the commercial interests, the existence of a feasible outcome in political economy terms, or the sufficiency of time, an optimistic note was sounded by some observers. It was pointed out that rounds do get done. To be sure, there is a lot of posturing; but, in the end, negotiators cut a deal on the basis of what is on the table. For example, the last three Rounds—Kennedy, Tokyo and Uruguay respectively—resulted in tariff cuts of about 33 percent, 33 percent, and 33 percent. How hard is it to guess what

the next number will be? Realistically, farmers will get their money, even after an ambitious, successful conclusion to the Round; it is a question of which "box" the money falls into and just how trade-distorting the support will be. Ultimately, the Doha Round may not be a "big" result based on *ex ante* expectations and hopes; but, it was suggested, when we look back, it will in fact be seen as a big result.

What was needed to move things forward was to connect the various elements of the negotiation—services, NAMA, agriculture and other elements—in order that the trade-offs could better be framed. In this regard, the Hong Kong Ministerial had established a useful common deadline of July 31st 2006 for progress on agriculture and NAMA and for the first real services offers. Further, it was noted, the plurilateral approach that has been adopted for services is happily also a sectoral approach; this can drive a constructive dynamic and lead to a different kind of negotiation based on sectoral specifics.

To summarize, in response to the question "Is the Round doable?", the discussion yielded an answer that might best be characterized as "The Round is not undoable."

Development: the Major Conundrum in the Round

Without a doubt, development has been the most contentious and ultimately confused aspect of the Round. Some saw this as a congenital defect in the framing of the Doha Development Agenda (DDA); according to this view, the Round had been misconstrued from day one.

Some of the confusion reflects the fact that it was, as one observer put it, "a masterpiece of constructive ambiguity".

However, it cannot be ignored that the choice in Dubai in November 2001 had been to have a development round or not to have a round at all. The Round was launched on the basis that it would provide a response to the North-South divide that had emerged from the outcome of the Uruguay Round—regardless of whether the emergence of this divide was an unintended consequence or a reflection of the power imbalance in the negotiations, a point on which views differed. The problem lies there-

fore not so much in the origins of the Round as in the lack of agreement as to what the Round was to deliver with respect to development.

Trade and Development: Clarifying the Link in the Doha Round

As one observer put it, the term "trade and development" is somewhat like a Rorschach test: everyone sees something different. To some it means agriculture, to others it means "aid for trade" (or capacity building), to others it is about preserving policy space, and to others it is about market access (supported as necessary by capacity building). That being said, observers from both the developed and developing worlds argued that considerable progress had been made in clarifying the ambiguities papered over at Doha. As one observer put it, the use of the development label for the Doha Round did create confusion but we are now almost out of those woods.

In this regard, it was argued that the development community has come a long way in recognizing that development is very complex and largely domestic. While trade liberalization is seen as an integral part of restructuring economies to take advantage of globalization, the idea of a trade round driving development, it was suggested, has the tail wagging the dog. The practical problems facing the would-be exporter in a typical developing country—e.g., long delays and theft of goods in transit—are not matters that the WTO can deal with. Similarly, WTO rights and obligations represent only a small part of the development function—for example, it was noted, it is not the handful of artists in developing countries who actually have foreign sales who need intellectual property protection but rather the thousands who work in the domestic market who do not.

Moreover, it remains difficult to bring development into the mercantilist negotiating framework of a trade round. For example, it was suggested that there was a fundamental lack of coherence in the framing of the negotiations: the development dimension of trade is identified with market access but market access also means own progress on policy reforms (since a tax on imports is a tax on exports) and the least developed countries

9

were to be given a free pass in this regard. It is an optimistic view, it was suggested, that sound development outcomes could nonetheless be achieved through greater cooperation between the World Bank and the WTO processes such as the discussion of policy frameworks supported by trade-related technical assistance (TRTA) in the context of the "integrated framework". The integrated framework, it was suggested, is not working. For one thing, it is hard to sort out when aid is just "aid" and when it is "aid for trade"—not to mention when it is "new" and when it is just "re-profiled" existing money (the sense of some observers is that there is actually little if any new money). It was pointed out that "aid for trade" would involve earmarking aid, which goes against the last five years of World Bank policy against such earmarking. And there is a confusion of programs—Aid for Trade, Poverty Reduction Strategy Papers (PRSPs), and Millenium Development Goals (MDGs) are all overlapping frameworks for disbursement of aid.

Given the current perspective, some observers found it hard to imagine that the notion of trade ministers delivering "development" was ever taken seriously. This would require, it was suggested, a level of coordination of domestic agencies that is difficult, and of international agencies that is impossible, to achieve. A multilateral trade agreement was not going to be the catalyst for such a coordination of effort.

In turn, it was argued, this more realistic perspective has brought a new-found clarity as to what development now means in the Round. Conceptually, it was argued, the term "development" in the Doha Round context must be recognized to mean development as understood by trade negotiators rather than by the development community. That is why development in the Round is associated so closely with agricultural trade; the WTO is the forum in which agricultural trade is best negotiated and that is the area where the trade negotiations have the greatest leverage on the development function—even if agriculture might have been "oversold" as a development tool. Thus, at a minimum, development in the Round means increased market access in agriculture, which is what was not obtained in the Uruguay Round and which continues to be lacking in the Doha round.

10

Further, the idea of "policy space", problematic as it may be to those who see a major benefit to a country from the acceptance of trade disciplines, is now accepted. The Hong Kong Declaration reflected agreement on buy-outs in the TRIMs agreement for local content, performance requirements and so forth. This provides practical policy space. It was argued that we need to "declare victory" in the WTO on this issue and get on with the real WTO agenda which is market access.

To be sure, it was acknowledged that market access, while necessary, is not sufficient. A distinction needs to be drawn between what might be termed "market access" and "market entry". Improved market access under trade rules (i.e., through tariff reduction) can be negotiated but this does not guarantee market entry which also depends on the ability of an exporter to comply with the various standards that products must meet (SPS, TBT, ISO, Walmart specifications, etc.). And the costs of compliance are high. Capacity to take advantage of market access concessions obtained in negotiations is thus also necessary.

However, it was argued that capacity building is really a sideshow for the WTO, notwithstanding the fact that it has become the centerpiece of the discussion of trade and development in the negotiations: the bottom line is that the WTO cannot provide funds for restructuring but can deliver market access, starting with agriculture but including non-agricultural goods and services. In this sense, the label "development" has become a liability for the Round, distracting attention from the central WTO agenda.

Risks to the Multilateral System from A Failed Round

Paradoxically, the reappraisal of the role of trade liberalization in development is taking place at a time when the more dynamic developing countries have gained a major stake in the multilateral system and the least developed countries (LDCs) have arguably taken over from the United States and the European Union as the custodians of the multilateral system. There was a time when the least developed (LDCs) feared the WTO. Now, it was suggested, the LDCs have bought into the notion of a rules-

based system and are happy to play the WTO game while the developed countries are avoiding engagement.

This situation highlights a problem in the political economy of the Round.

The great byproduct of the trade and development linkage, it was suggested, has been to force a more sophisticated examination of the distribution of global welfare gains from liberalization and of the development impacts of particular aspects of agreements. In the course of this examination, it has become apparent that LDCs face a real risk of no welfare gains from a deal due to the negative effects of preference erosion and higher food prices (the vast majority of LDCs are net food importers). Indeed, the perverse result of this situation is that the more engaged the LDC is in exports, the greater the welfare loss from deteriorating terms of trade.

At the same time, it was suggested that the huge beneficiaries of agricultural liberalization would actually be the industrialized countries, where food prices would be lower. But the ground work to prepare the case for agricultural liberalization based on the actual nature of the benefits has not been done.[2] Nor, it was suggested has adequate homework been done in the industrialized nations to explain to farmers what agricultural trade liberalization means and how government will support incomes—it is possible, it was pointed out, to be generous to farmers in non-distorting ways.

Meanwhile it was argued, the major emerging markets—China, India and Brazil—have not been making the contribution commensurate with the benefits that they have been getting from the system.

This combination of interests and engagement, it was felt by many at the Roundtable, did not augur a successful conclusion to the Round. The EU and the US will not provide the

---

[2] As a footnote to this discussion, it was noted that this is not a new situation: the Uruguay Round launch effort included a visit to Japan to demonstrate the benefit of low food prices. But it didn't work – however, this effort met with an argument that meat prices in Japan were high because Japanese consumed so little meat!

leadership to shore up the multilateral system but the emerging markets and the LDCs are not ready to take over leadership.

This is problematic since, in the consensus view, there seems to be no alternative to the multilateral system to address many key and still unresolved issues that continue to disturb the smooth functioning of the global trade system.

First, an agreement could help clarify systemic questions raised by the present lack of a coherent view of where the multilateral system is going. The consensus on embedded liberalism has collapsed. The old multilateral framework based on the GATT has evolved into something more than just a trade system; as one observer put it, it is at present a strange and unbalanced amalgam with the "rights" of one factor of production—intellectual property—being protected but not others. There is no consensus on this "system". It is driven by ad hoc, episodic liberalization, mingling foreign policy with economics—"episodic ad hocery" one might call it. At the same time, it was argued that the broader systemic issues could not be handled within the WTO negotiations alone. The ability to achieve consensus within a group of 149 Members[3] is limited; this constrains the areas on which the WTO will be able to move. Hence labour, investment, etc. cannot be built into the multilateral framework in any significant fashion. Other mechanisms are needed for these areas.

Second, without an agreement, the WTO would be weakened as an institution—it would be a ship without a rudder. The momentum for reform would be dissipated and the WTO's ongoing role as "overseer" of the multilateral system would be weakened.

Third, the dispute settlement mechanism in particular needs an agreement. Without the "legislative" guidance provided by a broad agreement amongst the Members, the evolution of the system would increasingly be based on decisions by the judicial arm through settlement of disputes. This raises new issues since, with the expiry of the "Peace Clause" in the Agreement

[3] Editor's note: the number of WTO Members has since increased to 150 with the accession of Vietnam on January 11, 2007.

13

on Agriculture, the subsidies agreement also applies to agriculture. The proliferation of RTAs is also affecting the scene for dispute resolution by providing, in some instances, for a choice of forum in which to pursue dispute settlement.

Fourth, if the Doha Round were not successful, the scene of action would shift increasingly to the arena of regional trade agreements (RTAs). However, it was noted, RTAs themselves are embedded in the multilateral system. The WTO is needed to provide some discipline over the formation and operation of RTAs, especially in respect of rules of origin (ROOs). Multilateral tariff elimination would of course clean up the ROOs mess. Moreover, some things cannot be done in bilateral agreements, in part because of free rider problems and in part because the big players cannot deal with each other in a bilateral context. For example, it was noted, the EU and Mercosur have found it difficult to deal with sugar because of the nature of that market. This drives RTAs towards small deals. Further, RTAs are not uniformly successful in leveraging greater trade (south-south RTAs appear to have had a weak track record) for those that can conclude agreements while raising a problem of exclusion and preference erosion for those that cannot. Finally, it was remarked that, if trade is about integration into global supply chains, RTAs are not helpful, they in fact are the opposite— they can restrict access.

In short, there is no perfect substitute for the WTO; some things will get done in the WTO or not at all. The cost of failure of the Round would be damage to the WTO's credibility which would represent an important system failure. The resulting drift would expose the system to developments, which could include disruptive change—including protectionist action against China, significant currency realignments and so forth.

Conclusion

With the window of opportunity for a timely successful conclusion to the Doha Round rapidly narrowing following the Hong Kong WTO Ministerial, close observers of the international trade scene underscored the need for a comprehensive agree-

ment to shore up the multilateral system, identified necessary next steps to bring about such an agreement and, notwithstanding numerous obstacles, pointed to reasons to believe that such an agreement could in fact be achieved. At the same time, the political economy of the Round was not seen as especially propitious for a successful conclusion; expectations were accordingly being trimmed (e.g., to a "Doha lite" outcome) and an extended hiatus was seen as a very real possibility, exposing the system to additional pressures and risks. Any optimism about a successful conclusion to the Round was thus qualified.

Events over the course of 2006 initially tended to validate the more pessimistic assessments of prospects for the Round but subsequently started to reflect the more hopeful views.

It's never over till its over.

Part II

A Conference on Integrative Trade

Integrative Trade: Issues for Trade Analysis, Statistics and Policy

Christopher Maule[*]

On December 6, 2006, the Centre for Trade Policy and Law at Carleton University organized a conference on the theme "Integrative Trade between Canada and the United States—Policy Implications". The discussion was structured around three main papers, which constitute the next three chapters of this volume:

- *Timothy Sturgeon,* "Conceptualizing Integrative Trade: The Global Value Chains Framework", *which provides an overview of the state of development of multi-disciplinary research on the evolution of global value chains;*

- *Art Ridgeway,* "Data Issues on Integrative Trade between Canada and the US: Measurement Issues for Supply Chains", *which examines the issues posed for statistical agencies in grappling with the changing international industrial landscape; and*

- *Michael Hart and William Dymond,* "Trade Theory, Trade Policy, and Cross-Border Integration", *which examined the implications of these developments for the trade policy community.*

This note sets out the background which motivated the organization of the conference and describes some of the concepts which shaped the discussions.

---

[*] Professor Emeritus, Department of Economics; c.maule@rogers.com. The views expressed in this note reflect the views of the author and are not to be attributed to Foreign Affairs and International Trade Canada or to the Government of Canada.

Introduction.

International commerce is being re-shaped by the fragmentation of the production process—the splitting up of the stages of production and locating them in different places in the global economy—and the resulting expansion of trade in intermediate goods and services and inward and outward investment. Glen Hodgson refers to this as integrative trade [1]

In discussing integrative trade, new terms have entered the lexicon of industry studies, such as outsourcing, offshoring and supply chain management. Previously, the concepts of just-in-time production and total quality management were introduced to the analysis of industries. Do these new terms point to new phenomena or do they simply represent new labels for familiar activities? What implications does integrative trade have for the statistical agencies charged with measuring international commerce? And what issues do these developments raise for the conduct of trade policy?

To set up the discussion of these issues in the next three Chapters, this note describes some of the terms which have come into increasingly general use in discussing modern international commerce, looks at how different disciplines address the ways in which industries are organized to see if they throw light on policy issues, and notes some of the issues associated with data sources and what further research might be needed.

Some Terminology

Outsourcing and Offshoring

In public discussion, outsourcing refers to situations where firms purchase inputs from other firms as opposed to producing the inputs themselves. It is further refined to distinguish between outsourcing that takes place at home and abroad. Foreign

---

[1] Glen Hodgson, "Trade in Evolution: the Emergence of Integrative Trade," EDC Economics, March 2004, p. 5. See also Hodgson, "Integrative Trade and the Canadian Experience," EDC Economics, May 2004, both accessed at www.edc.ca. Other references to concepts surrounding integrative trade are found in the conference papers.

outsourcing is referred to as offshoring. A further distinction is made between outsourcing from a plant owned by the firm making the purchase or from an independent firm. Thus there are four possible cases:

1. Firm in Country A purchases from one of its plants in A
2. Firm in A purchases from an independent firm in A
3. Firm in A purchases from one of its plants in Country B
4. Firm in A purchases from an independent firm in B

All four cases involve some degree of outsourcing but only Cases 3 and 4 are offshoring. Cases 1 and 3 involve intra-firm transactions and transfer pricing but only in Case 3 does transfer pricing involve international trade. Case 4 involves trade but not transfer pricing since the transaction is arms-length between buyer and seller.

In business terms, outsourcing is part of the ongoing evaluation that a firm's management makes about whether to make or buy inputs when determining the most cost efficient way to organise production. As conditions change in different markets including changes in technology the make-versus-buy decision is reviewed.

As examples, offshoring in manufacturing takes place when a Canadian shoe retailer purchases shoes made either by its subsidiary plant in India, or by an Indian owned manufacturer. In the resource sector, Alcan supplies its Canadian smelters with bauxite and alumina from abroad, either from its own plants or from independent suppliers. It has no option to offshoring for these inputs as there are no commercial deposits of bauxite in Canada, but what Canada has is the energy needed to convert alumina into ingot.

Offshoring in the service sector has received much of the public attention to date, in part because information technology has made it easier for firms to contract out for service activities abroad that were previously performed at home. Thus North American firms locate call centres, software programming activities and finance and accounting functions in countries such as India and the Caribbean. Technology has made it possible for many service activities to be much more footloose.

Public debate has focused on the employment impact of off-shoring which represents only a part of the economic impact. There is nothing new about the general process of offshoring except that technological change now means that it affects a different and wider set of firm functions. In the past textile jobs migrated from North America to Asia due to lower costs in Asia, today accounting and programming jobs are migrating out due to lower costs made possible by technological change. In resource-based activities, such as agriculture and minerals, the extent of outsourcing depends on where the raw material and its final market are geographically located. The statement by Greg Mankiw that "services offshoring is just another form of trade," reflects the situation that technology now allows offshoring for a new set of activities affecting different occupations than was previously the case for offshoring manufacturing and resource activities.

Supply Chains

Supply chains describe the stages of production organized by firms to manage its operations. Managers purchase inputs which are then converted through value-added stages of the production process into outputs for sale to other firms as intermediate goods or to final consumers as end products. For example, steel firms have a supply chain of inputs that leads to the production of steel for sale as intermediate goods to automotive and other firms. Automotive firms purchase steel and other inputs for manufacture and assembly of cars for sale to final consumers. Part of manufacturing production involves tangible goods but production also requires service activities such as R and D, design, planning, finance, advertising, labour relations, transportation, and storage. Each of the required inputs of goods and services can be subjected to the make-versus-buy supply chain decision that can influence whether or not it is outsourced and if outsourced whether it is off-shored. The stages of a typical firm's organization are shown in Diagram 1 (below) in terms of a firm's primary and support activities, each of which has the potential to be located in a particular place.

20

Diagram 1:
Value Chain of Firm's Primary and Support Activities
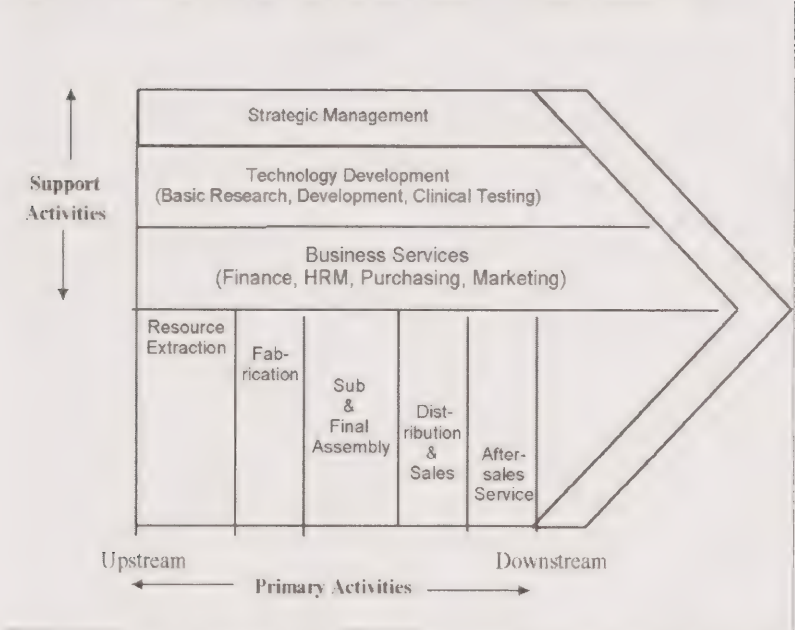
The supply chain or the successive stages of production for a firm will vary by firm and industry. The stages are sometimes referred to as the value chain and thus use is made of such terms as supply chain management or value (sometimes value-added) chain management. In economic literature on industrial organization the term vertical integration is used to describe the supply chain. Vertical integration refers to the extent to which stages of the production process are contained within a firm. Thus supply chain and vertical integration refer to similar aspects of a firm and industry. In economics there is an extensive literature that discusses aspects of vertical integration, while supply chain management is often used in discussions of business policy.

Just-in-Time and Total Quality Management

Two other terms used in management literature, just-in-time (JIT) production and total quality management (TQM), relate to aspects of the supply chain. JIT refers to the way in which goods move from one stage to another in the production process. If the item produced at Stage 1 is required as an input at Stage 2 of the production process, then Stage 1 output can be produced and stored ready for use at Stage 2 in which case there are inventory costs between the stages. Alternatively, the output at Stage 1 can be produced just in time for use at Stage 2, thereby reducing inventory costs. But organizing production in this way may give rise to other costs. Suppose there is an interruption in the supply of Stage 1 output, then, with JIT production, Stage 2 has to cease operating as there is no inventory to call on. The reduced costs of holding inventory can be offset by the risk and costs associated with production stoppages. Interruption can occur for numerous reasons such as defective parts, the breakdown of machinery or the failure of deliveries to arrive. If these deliveries involve cross-border shipments then customs clearance must occur.

TQM refers to the idea that if there is 100% checking of inputs for quality in production processes then the stages of production will flow more smoothly, reducing or eliminating the need for inventories of parts at each stage in the process. TQM supports the functioning of JIT and reduces costs of operating and managing the supply chain. TQM may involve higher costs of monitoring product quality but can reduce inventory costs.

Technology and the Supply Chain

Along with economic and population growth, technology has had an impact on the worldwide growth of international trade in a number of ways especially relating to transportation and communications. Lower transportation costs have occurred in maritime shipping through containerization, in railways with the use of high speed and unit trains, in pipelines used for liquids and solids, and in air transportation with jumbo jets and Fedex-type overnight delivery services.

In communications, digitization means that information of all types (print, audio and video) can be coded as digital signals and shipped by wired and wireless means. Activities that were previously undertaken within or close to a manufacturing plant can now be done more cheaply at a distance. Call centres are located in India and the Caribbean for companies located in North America and Europe; clerical services for insurance and financial companies are dispersed around the world; software programming, consulting, accounting and other service activities are now more easily traded. The services segment of the labour force that previously was partly protected from foreign competition now, because of technology, faces competition from cheaper labour in other countries.

Activities that were once considered as being non-tradable are now traded. Examples of non-tradable activities include services like haircuts, restaurant meals and funeral parlours where the supplier and customer have to be in the same place, but even in these cases a customer in one country can travel to the supplier in another to receive the service. The supply of medical tourism is precisely this with the patient traveling to medical facilities in another country and is reflected in the GATS Mode 2 form of service supply. While some service items may be difficult to trade, most can be subject to some form of trade and technology increases the likelihood of trade.

Supply chain changes over time

The term Fordist production refers to the early organization of vertically integrated automotive firms, a format initially copied by Toyota and other Asian car manufacturers. With Fordism, the design of an automobile, manufacture of parts, assembly, sales, financing and promotion were all conducted by the same firm. It might have many departments and divisions by functions and products for its primary and support activities but these were coordinated domestically and internationally by a head office. Such a firm could become multinational with some functions and products undertaken abroad. Many firms in natural resource and manufacturing developed in this way.

23

The same Fordist vertical integration was the case initially for firms in many service industries such as banking, advertising and management consulting. Over time, as a result of changing costs conditions including the effects of trade barriers and policies affecting the ownership of foreign direct investment, the cross-border movement of persons and the licensing of technology, these vertically related structures became modified and the supply chain altered as ways were found to reduce costs.

At the other end of the Fordist spectrum are companies that own few assets. They design products, arrange for their manufacture by others and organize delivery of the products to customers who also receive post-sales service arranged by the company. Dell Computer is one example—see Box 1 below. It designs computers based on parts such as Intel manufactured processors, hard disk drives and flat screens produced and assembled elsewhere. Dell receives orders from customers that it transmits to its manufacturers who in turn order the components needed to assemble the computers. It then arranges for shipment and for the provision of call centres to support customers and provide for warranty service. Production and inventory management are activities that tie up capital and expose the company to financial risk as sales fluctuate. By performing only a few of the vertically related functions and merely coordinating the others, Dell reduces but does not eliminate its exposure to risk. It has few of the direct costs of production but it depends on the reliability of suppliers for the quality and timely delivery of their products. While some costs are reduced there exists the potential for others to increase. WalMart and Ikea are other examples of firms that follow this model for some of the products they sell to consumers. Kenney and Florida (2004) provide other industry examples of firm locational decisions.

Box 1: In 2005, a Dell laptop was designed in Texas and assembled in China. It had its keyboard made in China, the motherboard in Malaysia, the flat screen in South Korea, and the software was compiled in the US, India, Sweden and Russia. The product label stated "Made in China," although establishing nationality seems to be an arbitrary process (Gave, 2005:10)

In order for a more vertically disintegrated or fragmented approach to industrial organization to work, an efficient communications and transportation infrastructure is needed as well as the absence of government policies that impede the cross-border movement of trade in goods and services. There also has to be a level of trust between buyer and seller and confidence in the judicial system to settle fairly any disputes that arise.

These examples illustrate how technology permits different ways of producing, distributing and coordinating industrial activity, that is different ways of organizing an industry's supply chain. As change occurs, some goods and services disappear or diminish in importance such as the typewriter, black and white television set, and postal services, while others increase such as email communications and the transfer of digital files of audio, video and print materials. For example, copies of print encyclopaedias still exist but new ones are hard to buy, while online versions such as Wikipedia are competing with print versions. Each firm examines it supply chain to see how technology can be introduced to reduce costs and make the firm more competitive.

Offshoring and supply chain management is also related to policies affecting the movement of labour. If North American firms outsource clothing manufacture to Latin America and China because of lower labour costs, one alternative would be to import labour to North America to do the work. This happens and is reflected in the temporary work visas given to migrant workers in North America as well as the inflow of illegal workers; the estimated number of illegal workers in the US is currently 11 million. In some instances the configuration of the industry supply chain allows work to be sent to the workers as in the case of call centres, while in others the workers have to come to the work as in the case of harvesting agricultural products where mechanization may not provide as efficient an alternative.

Disciplinary Approaches to Industry Studies

Literature that examines changes in industrial organization can be found in economics, business administration, geography and sociology as well as in discussions of trade policy. Each disci-

pline has its reasons for making such a study and a particular framework and terminology for its analysis. Students of economics, business administration, geography, and sociology study industries, but through different lenses and for different reasons. Each examines a series of issues some of which are overlapping and some unique to the discipline's focus. One common denominator is industrial organization, but that term is used for different purposes. For example, economics stresses issues of competitiveness and efficiency, business administration the overall performance of firms, geography the location of production, and sociology issues such as industrial development and poverty alleviation. Before exploring these disciplinary differences, we outline what is meant by some frequently used terms.

Economics

The field of microeconomics contains the subfield of industrial organization that examines the way in which firms can be grouped into industries in order to assess the extent of competition and the consequences for society of competitive conditions, for example when markets exhibit different degrees of monopoly power with the ability to influence prices[2]. A main concern of economics is the efficiency of resource use. Its focus on market power is because less competitive markets are likely to result in a waste of resources from the viewpoint of society as a whole. Among the main factors considered in industry studies are:

- The importance of economies of scale influencing the extent of horizontal integration.
- The extent to which firms are vertically integrated thereby owning stages in the production process, for example in the oil industry the stages of exploration, production,

---

[2] Industrial organization in economics has two dimensions, the organization of firms within industries and markets which leads to a focus on competition, and the organization of the firms themselves by functions, by divisions and departments and as domestic and multinational firms. The two are related in that the efficiency with which a firm is organized will affect its ability to compete in a market.

transportation, refining and distribution of refined products for sale to final consumers.

- The degree of product differentiation examines whether a firm produces one type of product, such as small cars, or a range of products, such as small, medium and large cars, trucks and buses. The extent of competition with other producers can vary depending on the range of products produced by each producer. This dimension is also referred to as diversification, but diversification can be both by product and by regional location. Note, vertical integration can also be considered a form of diversification by stage of production as opposed to by end product. The concept of scope economies is used to help explain the extent of diversification within a firm.

- The ease with which new firms can enter a market where the firms may be either newly established or firms in another industry that diversify into a new industry; and the ease with which failing and other firms can leave the industry.

Industrial organization is not especially concerned with where a firm locates the various stages of its production process, an interest that falls more within the economic fields of international trade and investment. A combination of cost factors associated with the above four sets of economic conditions and the effect of political boundaries will be major influences on where activities are located geographically.

Note that economics distinguishes between an industry and a market—see Box 2. The cement industry in Canada may be represented in all regions of the country but it consists of separate markets. For example, for reasons of transportation costs, cement producers in Atlantic Canada and British Columbia will find it difficult to compete in each other's markets. This is an example where scale economies in production may be offset by freight costs of reaching consumers leading to more and smaller plants than might exist in a market with higher population density. Distance becomes an important factor in how a firm organizes its activities.

Business Studies

Faculties of business share many interests with economics but business studies focus more on factors concerned with managing the firm, thus they offer courses on topics such as production, finance, marketing, strategic planning, organizational behaviour, and advertising and promotion. Their interest is in the factors leading to firm success measured in terms of return on investment regardless of whether this is associated with competitive or monopolistic market conditions.

Discussion of strategic and tactical decision-making and planning is an important concern of business studies and leads to examination of the organizational structure of the firm in terms of factors such as make-versus-buy, outsourcing, offshoring and the extent of vertical integration and diversification. All these terms relate to the idea of supply management.

Geography

In contrast with the concerns of economics on questions of efficiency and competitiveness, and of business studies with reasons for firm performance, geographers are often more interested in explaining the location of industrial activities. They look at the same set of activities but with a different focus. The economics of agglomeration is used to explain why particular industries cluster in certain areas, for example film and television production in Hollywood and Mumbai, financial services in London, New York and Tokyo, hard disk drive assembly in Singapore, semiconductors in Silicon Valley, and the manufacture of clothing in China and Mexico.

The location of natural resource and agricultural production are strongly influenced by resource endowment and climate, but manufacturing and service activities often require a different explanation. Any country (location) could have a film industry, and many do, but production tends to be concentrated in certain places. The US, India, Egypt, Hong Kong and Nigeria are examples of locations where a significant amount of film production occurs, but the technology required to make films is widely available.

Geographers look at the process of agglomeration or the benefits of grouping similar or related activities in the same location. Film production requires a wide range of support activities such as sets, props, carpenters, technicians, costumes, hair dressers and makeup persons, stunt performers, animals, scenery, composers, and musicians, in addition to producers, directors and performers. Once production in a location reaches a critical size these inputs are attracted to the production site and become available for a number of film companies. The

29

benefits of locating in one place provide reasons for the structure of industry organization that occurs. The contribution of geography to explaining industry organization is similar and complementary to the approach taken by economics.

Like other disciplines, geographers focus on the supply chain or the vertically related stages of an industry's production process but with emphasis on where each of these stages is located, as opposed to the economist's concern, for example, with efficiency and competitiveness. The concepts of vertical integration and supply chain are similar if not identical although they tend to be used for different purposes by the different disciplines.

Sociology

Work by Garry Gereffi and others have focused on the organization of industries (Gereffi, 2005:79)

> "...the starting point for understanding the changing nature of international trade and industrial organization is contained in the notion of a value chain, as developed by international business scholars who have focused on the strategies of both firms and countries in the global economy. In its most basic form, a value-added chain is 'the process by which technology is combined with material and labour inputs, and then processed inputs are assembled, marketed, and distributed. A single firm may consist of only one link in this process, or it may be extensively vertically integrated...'(Kogut, 1985:15).

Concepts used in this analysis that overlap with other disciplines are the significance of transaction costs, the vertical disintegration of multinational corporations, core competencies of firms, the growing international trade in components and intermediate products as opposed to final goods and services, information flows and the variation in value chain governance from markets to hierarchies.

In contrast to the efficiency concerns of economists and the locational interests of geographers, sociologists have emphasized factors that affect "... not only the fortunes of firms and the structure of industries, but also how and why countries ad-

vance—or fail to advance—in the global economy." Their interest is in crafting "...effective policy tools related to industrial upgrading, economic development, employment creation, and poverty alleviation. (Gereffi, 2005: 79). The aim is to explain how particular industries can assist in a country's economic development. The extent to which stages in a production process can be located regionally provide opportunities for the dispersion of production in different parts of the world and to countries at different stages of economic development.

Other Disciplines

Disciplines such as law and political science also have an interest in industrial organization. Corporations are the principal form of organization for industrial activities. Corporate and contract law are vital to an understanding of how firms operate and do business with each other as are the branches of law dealing with bankruptcy, taxation, labour, the environment, trade, investment and intellectual property. Politics is concerned with the concept of power. Corporations are able to exercise power in numerous ways thereby affecting the sovereignty of states, a topic examined by political scientists.

Outsourcing, Offshoring and Supply Chains: Data Issues

Public discussion of outsourcing stems mainly from its employment impact that now, due to technological change, affects services as well as goods. Debate is reinforced by statements that industrialized economies are primarily service economies with around 70% of employment being services related, while declining employment shares are associated with the manufacturing and resource sectors. What does this mean and how is it measured?

    The distinction between goods and services is enthroned in public debate, in data on production, employment and trade and in the WTO with the GATT disciplines for goods and GATS for services. Underlying the distinction are some difficulties. The production of a good, such as an automobile or a pair of shoes,

31

involves the production of tangible objects, while production involving financial, medical and engineering advice involves an intangible output. The production of music illustrates a combination of good and service. When transmitted over the air, a song is a service; when embodied in a disk, it becomes a good. Should music then be considered the production of a good or service or both?

It is not clear that the automobile and shoe examples are pure goods. If finance, advertising, transportation and warehousing are service activities, all are associated with and can take place within a goods producing firm. If the firm is classified as goods producing it is because most of its value added activity is associated with the tangible side of the product and not with the intangible service inputs. Assume that the firm decides to contract out for some or all of its financial, advertising, transportation and warehousing requirements thereby reducing its work force, the final output of the firm may remain the same but the configuration of inputs used to produce the output has changed. The firm now out-sources for these services which when counted separately make it appear that the service sector has expanded at least in terms of employment.

What has actually happened is that the services once performed within the manufacturing firm are now contracted out with a contraction of manufacturing employment and an expansion of services employment. Management is continually engaged in reconfiguring the firm's supply chain in order to reduce costs and remain competitive. In so doing the national economy may appear to become more service oriented when in fact little has changed. Data on employment by occupations should provide more accurate information of what has actually changed. It may well be the case that developments in information technology have led to the need for more persons to be engaged in the provision of programming services, the operation of call centres and the supply of repair services, in which case the occupational structure of an economy's labour force may change with a greater emphasis on services. Some combination of the needs of new service industries and the reconfiguration of production within existing goods producing industries probably ac-

counts for the larger percentage of persons employed in services, and the finding that economies are becoming more service oriented. Data should be able to provide a more precise description and explanation of these. What do the data show?

The answer is not clear but the work needed to provide an answer is set out in a recent report published by the Industrial Performance Centre at MIT.[3] It concludes that in the case of the US, there is an absence of adequate data on services traded internationally; a similar absence of data on domestic trade in services; and a lack of adequate data on employment by occupation and industry. Until such data become available, it will be difficult to assess accurately the extent of outsourcing and offshoring by industry and the trade and employment impacts. The MIT study notes that the classification system for traded items includes 16,000 categories for goods versus only 17 for services. Similar discrepancies exist for goods and services traded domestically within the US and for the occupations associated with goods and services production. Goods trade has always been easier to measure because a tangible object is involved for which customs paperwork is required before it can cross a border to its destination as an import. Exports usually have no such requirement and their record may depend on surveys taken in the seller's country or import data supplied by the buyer's country. For example, for goods exports to each other, the US and Canada rely mainly on import data from the other country. Measurement of services trade depends largely on surveys conducted in each country with all the problems associated with ensuring that the surveys are completed[4].

The under-reporting of the service sector is tied up in part with the underground or informal economy whose size varies by country but even in the case of a developed economy such as Greece is large. A recent report in the *Financial Times* (Sept.

[3] "Services Offshoring Working Group, Final Report," lead Author T. J. Sturgeon, September 10, 2006, MIT Industrial Performance Center http. web.mit.edu ipc publications pdf IPC_Offshoring_Report.pdf.

[4] In the US the threshold for collecting services trade data has been $6m per annum for imports and $8m per annum for exports.

29, 2006) noted that "Greece suddenly found itself 25 per cent richer on Thursday after a surprise upward revision of its gross domestic product, the fruit of a change to national accounts designed to capture better a fast-growing service sector including parts of the black economy such as prostitution and money laundering."

With services, there is considerable domestic and international trade conducted for no charge or trade that may be under-reported. Users of Skype make telephone calls in Canada and abroad for no charge. Foreign published newspapers and journals are read for free on the Internet. International calls are made by using a phone card whose value is probably not captured in services trade data. These are examples where border policies have little or no impact on trade. For trade in goods, borders are a much more significant issue.

Conclusions

A new and complex international commercial landscape is emerging, which is spawning new concepts and new terminology, creating new demand for improved statistical measurement, and raising questions for traditional trade policy and practice. It is timely that these issues be aired in Ottawa, as Canada seeks to consolidate its place in the North American production platform and enhance its role in the global division of labour.

References

Gave, C., K. Kaletsky, and LV. Gave. 2005. *Our Brave New World.* Hong Kong, GaveKal Research.

Gereffi, G., J. Humphrey, and T. Sturgeon. 2005. "The governance of global value chains," *Review of International Political Economy*, 12(1): 78104.

Kenney, Martin with Richard Florida. 2004. *Locating Global Advantage: Industry Dynamics in the International Economy*, Stanford University Press.

Kogut, B. 1985. "Designing global Strategies: Comparative and Competitive Value-Added Chains", *Sloan Management Review*, 26(4): 1528.

Conceptualizing Integrative Trade: The Global Value Chains Framework

Timothy J. Sturgeon[*]

Introduction

The global economy has entered a new phase of deeper, more immediate integration that is exposing national and local economies to the winds of global competition as never before. These winds can fill the sails of our domestic firms and industries, or blow them away. Peter Dicken (1992: 5) has argued that an earlier era of 'internationalization,' characterized by the simple geographic spread of economic activities across national boundaries, is giving way to an era of 'globalization,' which involves the *functional integration* of these internationally dispersed activities. It is this functional integration that drives our growing interest in 'integrative trade.'

What is it that enables greater functional integration in the global economy? Two key differences with the past are rapidly increasing industrial capabilities in developing countries, capabilities that reside both in local firms and the affiliates of multinational firms, and new computer-mediated approaches to real-time integration of distant activities. These new features facilitate international trade in many intermediate goods and services that have not previously been sent across borders. As a result, opportunities have opened up for firms to engage with the

[*] Senior Research Affiliate, Industrial Performance Center, Massachusetts Institute of Technology, and COE Research Fellow, Institute for Technology, Enterprise, and Competitiveness, Doshisha School of Management, Kyoto, Japan. This paper was prepared for the conference, Integrative Trade between Canada and the United States — Policy Implications, organized by the Centre for Trade Policy and Law, Ottawa, December 6, 2006. The views expressed are those of the author and not to be attributed to the Department of Foreign Affairs and International Trade.

global economy—as buyers, suppliers, sellers, distributors, contractors, and service providers—in ways that were impossible even a few years ago. However, recent changes have created new challenges and risks, as well as opportunities. The growth of integrative trade has served to expand the arena of competition beyond final products to the vertical segments and business functions within and across industries, raising the performance requirements for firms that may have been more insulated from global competition in the past.

The implications for policy are profound. How can workers, firms, and industries be provided with the best environment for engaging with the global economy? How can we be sure that enough wealth, employment, and innovative capacity are generated at home? These are open questions. Even if policy-makers seek no interventions in the areas of trade or industrial policy, global integration can make the process of economic adjustment more difficult because it accelerates the pace of change. Trade adjustment, education, taxation, innovation, and infrastructure are all policy areas in which we need to consider the effects of integrative trade. The problem lies in our shallow understanding of the process and of its specific effects.

In this paper I argue that global integration signals an urgent need to develop new conceptual tools. Effective policy responses require a clear, detailed, and timely view of global integration and related economic changes based on solid economic data. Good economic data also provide appropriate market signals for companies, workers, students, and educational institutions. Current economic statistics, at both the level of countries and globally, are clearly not up to the job (see Sturgeon et al., 2006). But we need more than more and better economic data. We need to restructure our thinking about mechanisms and outcomes in the global economy.

This paper is organized as follows. First, I outline five trends that are working to accelerate the pace of global integration. Then, I argue that an understanding of the core dynamics of global integration requires deep qualitative knowledge of the details of specific industries. Nevertheless, the results of such industry-specific qualitative research will remain of limited util-

ity unless they can be generalized in a way that renders the insights they provide industry-independent. As a way to begin to fill this gap, I offer the "Global Value Chains (GVC) Framework," an industry-independent conceptual model that highlights two critical aspects of integrative trade that are not captured by current economic statistics: *power and coordination* in the global economy. After presenting the features of the GVC framework, I go on to focus in more detail on the newest and most dynamic form of GVC governance: value chain modularity.

Five Trends Driving the Pace of Global Integration

Global integration has a variety of indicators. First and foremost is a boom in international trade. For example, in 1985 the value of China's exports to the United States stood at about US$6.5 billion, less than one percent of total U.S. imports, and trade between the two countries was roughly balanced. In 2005 the United States had a trade deficit with China of more than US$185 billion, and China's share of total imports had increased to more than eleven percent. Intermediate goods trade is rising faster than final goods trade, a trend that indicates that increasing specialization and value chain fragmentation are key features of global integration (Yeats, 2001; Feenstra, 1998; Hummels et al., 2001). Another indicator of deepening integration in the global economy is rising anxiety about the loss of white-collar jobs in North America and Europe, triggered in part by India's dramatic and very recent successes in exporting software and business services (Sturgeon et al., 2006). Even in fresh food, patterns of production and trade are geographically extensive, complex, and dynamic. Fruit, vegetables, meat, and fish are grown and processed around the world and delivered daily to supermarket shelves in Europe and North America (Dolan and Humphrey, 2000, 2001).

These quick examples suggest that change in the global economy is broad-based, and is proceeding with great rapidity. What are the drivers? There are five trends that are combining with increasing trade liberalization to accelerate the pace of global integration:

1. *The "great doubling" of the global workforce.* The end of the cold war and abandonment of autarkic "import substituting" development policies in places like India, Russia, and China have quite suddenly increased the size of the global workforce from approximately 1.5 billion to 2.9 billion (Freeman, 2005). If the energy and talent of these workers can indeed be effectively tapped (see below), this increase could prove large and sudden enough to place downward pressure on wages in both advanced industrial economies such as the United States and Canada as well as developing places that have long been part of the global system such as Latin America and South East Asia.

2. *Lower costs and greater capacity in global communications networks.* The overbuilding of international data transmission networks during the "dot.com" boom, as well as aggressive efforts by countries such as India and China to improve their international links and domestic infrastructure, have contributed to a radical and sudden lowering of the costs of tapping the workers and industrial capabilities that reside in developing countries. This has improved access to the huge pools of low cost but adequately skilled labour that have recently become available in the global economy.

3. *The standardization, formalization, and digitization of work.* There has been broad application of information technology to a wide variety of work tasks and business processes (e.g., word processing, call routing, inventory management, factory production). Information technology facilitates both the fragmentation and relocation of work and the reintegration of those fragments once tasks are completed (Bardhan and Kroll, 2003; Berger at al, 2005). As more firms have adopted information technology it has become more standardized to facilitate system inter-operation and information sharing (Levy and Murnane, 2004). The encapsulation of work tasks into standardized modules (Baldwin and Clark, 2000) eases the movement of work because it reduces the need for exchanging tacit knowledge and the amount of training or new capital investment required. Such "modular-

ity" is now quite common in manufacturing, but advancements are proceeding with great speed in services, in part because of what has been learned in the realm of manufacturing (Gereffi et al., 2005; Berger et al., 2005).

4. *The new, global supply-base.* Standardization has also helped to create new business opportunities for 'global supplier' firms that pool capacity for a range of customers (see Sturgeon (2002) for examples from electronics manufacturing and Batt et al. (2005) for call centers). Some of these suppliers are located offshore (e.g., in India, Canada, and Ireland), and others have become global in scope, with facilities in both advanced and developing countries. Such global suppliers specialize in collecting work from other firms and moving it to its "optimal" location on the globe. They make it easier for medium-sized and even small firms to engage in global sourcing and to locate parts of their business offshore.

5. *The rise of the global start-up.* Because of the above four trends, it has become possible for start-up firms to set up global operations from the first day of operation (Breznitz, forthcoming). Venture capitalists, in fact, are encouraging this practice (Wilson, 2003; Mieszcowski, 2003; and Grimes, 2004). Not only does this raise the possibility that a larger share of employment creation from new firm and industry formation will occur offshore, it also raises questions about the continued innovative leadership in advanced economies, since parts of the innovation process itself are being moved to developing countries. In industries such as electronic hardware, for example, firms based in North America have been able to retain (and in some cases regain) control over the innovative trajectory of some product categories while moving high volume, labour-intensive, and price-sensitive segments of the value chain to low-cost geographical locations. The question is how sustainable this is, and if a similar pattern can or will emerge in industries that are just beginning to become globally integrated, such as services?

The pace of change has emerged as a critical factor in the recent debate over the effects of global integration on advanced economies (Bardhan and Kroll, 2003; Blinder, 2005). There are three basic positions regarding how deepening integrative trade will affect developed economies such as the United States and Canada. These are spelled out in very rough terms as follows: (1) Specialization and innovative leadership will continue to make developed economies rich, so no policy interventions will be required (Bhagwati, 2004). (2) Policy-makers only need to worry if developed economies hive off parts of industries in which they have comparative advantages, but these negative effects will likely be small, so all that policy should aim to do is to compensate losers (Samuelson, 2004). (3) It is entirely possible for developed economies to lose comparative advantages over time, so policy-makers should take steps in some instances to assist existing industries and bolster innovative capabilities (Gomory and Baumol, 2000).

All of these positions suggest that time is required for successful adjustment to global integration. Innovation and new market creation take time to occur, compensating losers is only possible if there are not too many coming on stream too quickly, and the erosion of established comparative advantages might be staunched through policy interventions as long as it happens gradually. If change occurs with extreme rapidity, it will be difficult to innovate fast enough, to compensate the flood of losers quickly enough, or to craft and implement effective policy measures in time to make a difference.

The debate over the newest feature of global integration, services offshoring, is a case in point. Dossani and Kenney (2004; 32) argue that, in the realm of services, low capital intensity and the purely electronic form in which many services can be delivered will drive global integration faster than has been the case in manufacturing. And because service occupations are widely distributed throughout the economy, the negative effects of services offshoring could be more broadly based than has been the case with the offshoring of manufacturing work (Bardhan and Kroll, 2003). It may be that the flow of work offshore will be sufficiently large and rapid to make adjustment extremely difficult. In

this view, it is not that the theory of comparative advantage is wrong, but "…sometimes quantitative change is so large that it brings about qualitative change" (Blinder, 2005, p. 2).

The Importance of Qualitative Industry Case Studies

Because the stakes are so high, we must take global integration seriously and develop ways of thinking that place the new and emergent features of the global economy in the foreground. The venerable intellectual approaches to such questions focus on the roles of comparative advantage and transnational corporations in motivating and structuring international trade and investment. While these concepts have proved to be extremely robust and are still valuable, they do not emphasize the fragmentation of the value chain or the fluid, real-time integration of capabilities in advanced economies with capabilities in places that were all but outside of the global economy only two decades ago, such as China, India, Russia, and Vietnam. In fact, they emphasize the opposite: national export specialization in final products and the extension of national advantage, via multinational affiliates, to places without the domestic capabilities to effectively compete.

We should be concerned that the assumptions embedded in theories of comparative advantage may blind us to the truer nature of global integration: that industries are becoming globally distributed and are co-evolving in elaborate and ever shifting ecosystems that make it unclear where advantage truly lies. If we are to begin with a fresh sheet of paper, where should we start? One way is to move beyond aggregate statistics to work with microeconomic data ("micro-data") collected by government agencies. Over the past decade there has been a burgeoning body of research that relies on government-collected micro-data. Some of these resources have only recently become available.

There is a host of government programs that collect detailed economic data. Some of these programs, such as the Economic Census, use surveys to collect data for publication. Typically there are more detailed micro-data that underlie the published data. The mailing lists for these surveys can also contain valuable data on the basic characteristics of individual firms and

establishments. Other programs collect data for the purpose of administering government programs such as tax collection, compliance with environmental protection laws, and the like. For this reason such data are typically referred to as "administrative data."

One example of how administrative micro-data have been made useful for researchers is the US Census Bureau's Business Register, which is essentially the sampling frame for the Economic Census. Data included are business name, address, a unique establishment-level identifier, industry, employment, and the identity of the firm that owns the enterprise. Data about ownership allow the enterprises in the Business Register to be aggregated to the firm level. Jarmin and Miranda (2002) have assembled the Business Register into a time-series for 1976-2002, referred to as the Longitudinal Business Database (LBD). The potential of the LBD has just begun to be tapped. For example, Bernard, Jensen, and Schott (2005b) link the LBD to the universe of import and export transactions for 1993-2000, revealing a detailed picture of the characteristics of firms that do and do not trade and offering a wealth of research possibilities on how US firms' trading activities and domestic operations are related.

Another example from the United States is the Longitudinal Research Database (LRD), which contains data on all manufacturing establishments that were in at least one US Census of Manufactures since 1963 or one annual survey of manufactures since 1972. For 1992, the LRD incorporated data for over 378,000 manufacturing establishments (in non-census years the total is about one-sixth that amount). The LRD contains data that identify individual establishments, and a high level of detail on the manufactured inputs and products (outputs) of those establishments. Identification data include permanent plant and establishment numbers, industry codes, location, current status, and legal form of organization. Input data include total employment, number of production workers, hours worked, labour costs, materials costs, materials consumed, services and energy consumed, inventory levels, depreciable assets, and capital expenditures. Product data include receipts (value of shipments,

value added, value of re-sales); production details (5- or 7-digit SIC product codes, quantities of production, value and quantity of products shipped, value and quantity of interplant transfers, and internal consumption); and exports. Research using the LRD and other micro-data resources has explored a number of issues related to global integration, including establishment dynamics, job turnover, the effects of international trade, and productivity growth. While very valuable, these studies typically study the entire manufacturing sector and have not yet delved into the dynamics present in particular industries.

Researchers have also creatively used micro-data from more limited data sets to explore specific questions related to global integration. Harrison and McMillian (2006) and others have used the parent and foreign affiliate micro-data from the Bureau of Economic Analysis surveys on multinational firms to examine the relationship between affiliate activity and US employment. Swenson (2005) has examined the permanency of offshore assembly arrangements using extremely detailed data from United States International Trade Commission (USITC) reports. Kletzer (2002) has used micro-data from the Displaced Worker Survey to explore the experiences of workers displaced from manufacturing industries associated with increased foreign competition, and has made policy recommendations based on her findings. These studies are examples of leading-edge quantitative research on the employment effects of globalization. Because of the paucity of data collected on international trade in services, however, it is problematic to extend the methods used by these researchers to services.

But even micro-data are not enough. The rise in intermediate goods trade strongly suggests that we have moved beyond a situation where countries use domestic resources to develop and export products to the rest of the world. Countries and regions within countries are not responsible for making products and delivering services in their entirety, but have come to specialize in particular elements within the larger chain of value-added activities. As a result industrial output and export statistics provide a very partial view of where in the global economy value is created and where it is captured. Specifically, they provide very

43

little, if any, insight into the critical questions of how much control firms and industries in specific places exert over the activities they and others carry out in the global economy and how this control is translated into the distribution of gains among firms, countries, and communities. Because the picture of global integration provided by trade and investment data is so incomplete, the causal links to welfare indicators such as employment and wages derived from macro statistics can be weak and unconvincing.

What is required is deep knowledge of the forces driving change in specific industries, occupations, and geographic locations. Even with better quantitative information coming from analysis of micro-data, the impact of global integration on advanced economies will be extremely difficult to fully comprehend or respond to without a detailed view of how global integration is intertwined with other aspects of economic change, especially the automation and computerization of work and the prevailing characteristics of labour markets and corporate strategies in specific service industries and occupations. The best way to learn about the interaction of these complex elements of economic change is though qualitative research on the trade-offs that managers of individual firms and establishments in specific industries face and the choices they make. Ralph Gomory has referred to industry studies of this kind as "observational science."[1]

Over the past 20 years, grounded, qualitative, field-based research on specific industries has led investigators to a common set of questions and concerns. As industry after industry has developed deep connections beyond local and national jurisdictions, the practitioners of such "industry studies" have gravitated toward questions about how the global-scale division of labour is evolving, what specific roles firms based in different societies play in global-scale production networks, and what the implications of these differences are for the welfare and

---

[1]   Ralph Gomory is President of the Alfred P. Sloan Foundation. This comment was made in the course of remarks given at the Industry Studies Annual Conference in Cambridge, Massachusetts on Dec. 15, 2005.

economic performance of nations, workers, and communities, whose prospects and experiences are inherently more territorially bounded. In the 1960s through the 1980s, the multinational firm embodied the growing disjuncture between the motives of large firms and local communities. The concern was that the rise of "stateless" multinationals meant the demise of national industries and a loss of local control. But close observation shows us that even the largest firms remain rooted in their home economies in important ways even as their operations become global in scope. We are in the midst of a profound transition nonetheless. Multinational firms have arisen in many countries, resulting in a deep interpenetration of the global economy, driven by both outward and inward investment. But it is the expansion of non-equity ties, often referred to as 'global sourcing,' that generates the most novel and complex aspects of global integration.

Decisions about global sourcing and relocating business activities are inevitably made in the context of broader company strategies related to the development of new products, the pursuit of new customers and markets, the adoption of new technologies and production techniques, and the like. Distinguishing economic changes due to offshoring that displaces domestic employment from offshoring that does not—for example, when a firm establishes a presence to gain access to a foreign market that cannot be accessed through exports—is therefore extremely difficult to do without speaking directly with the managers making the key decisions. Even when examining the operations of a single firm, with full cooperation from management, it can be extremely difficult, if not impossible, to precisely measure the employment effects of global integration.

For example, Dossani and Kenney (2005), in their case study of Company X, an electronic equipment and services firm with approximately 30,000 employees worldwide, showed that the geographic consolidation of service-related activities in India was accompanied by simultaneous consolidation of business functions and information technology platforms (see Table 1). In the words of Rafiq Dossani:

Company X took the opportunity of preparing to outsource to India to completely re-engineer the way they did their back office work. In the process of doing this they created new job descriptions and new jobs in-house, new jobs for their local outsourcing partners, and new jobs for their offshore affiliates and partners. We tried to take a very granular view, to look at job descriptions, and follow where the work was being done, but found that this was impossible to do. So, even though we had an insider to work with and full cooperation, we were unable to actually look at job content and where that content was moved. For example, if a job consists of making an entry into a computer, and now it is made on a different platform, routed differently, supervised differently, it is not the same activity any longer.[2]

Table 1. The Context for Offshoring at Company X: Functional, Technological, and Geographic Consolidation

1) The consolidation of shared services across geographies and departments, particularly human resources, finance, engineering services and procurement, into a limited number of global hubs.
2) The consolidation of enterprise resource planning and customer relationship management [IT] systems into common platforms using off-the-shelf technologies and minimizing the usage of legacy applications.
3) The consolidation of geographical footprints.

Source: Dossani and Kenney, 2005, p. 25.

These methodological challenges should not lead us to abandon our efforts to gauge the employment effects of global integration, only to temper our confidence in estimates based on aggregate data or in the insights gained through qualitative research. Nevertheless, in specific industries and occupations, qualitative research can provide valuable insights into the real and potential job effects of global integration.

For example, Levy and Goelman (2005) use qualitative methods to show that only a tiny number of US radiology im-

---

[2] Author interview with Rafiq Dossani, February 2, 2005, Stanford, CA.

ages are currently read outside of the United States; they convincingly argue that it is highly unlikely that this number will increase substantially in the future. The shift from analog to digital radiology imaging has certainly made the remote analysis of radiology images technically feasible, a fact that has spurred much hand wringing in the media about radiology jobs "moving" offshore. Tight labour markets and high salaries for radiologists, in part due to a cap on federal funding for hospital residencies, also suggest high potential for the offshore interpretation of radiology images. But because there is a need, in many cases, for close consultation between radiologist and doctors, almost all radiology images are read at or very near the site where they are taken. Moreover, the high cost of radiology imaging equipment relative to the cost of interpretation, the restriction of US malpractice insurance to doctors who have done US residencies and passed US medical board exams, the group power of US doctors to restrict competition, and Medicare reimbursement regulations all work to keep the remote interpretation of radiology images on shore.

Because of these "institutional" factors, Levy and Goelman found that virtually all of the very small number of radiology images that are read offshore are read by radiologists who completed their residency and passed their board certification in the United States. For example, a US board certified radiologist in Sydney, Australia, can work days reading images generated at night in the United States. An understanding of such industry-specific factors, and their interaction, requires deep knowledge of specific industries and occupations that can only be gained through qualitative research methods.

As these examples show, industry case studies have the potential to reveal some of the deeper dynamics, and limits, of global integration. One core finding from this research is that firms from advanced industrial countries have played a central role in driving and shaping global integration. In India, firms that provide IT services interact with clients from around the world on a daily or even hourly basis to provide them with the packages of services they need (Dossani and Kenney, 2003). In horticulture, large retailers have worked closely with exporting

companies in Africa and Central America to obtain products that meet their ever-increasing demands for variety, food safety and speedy delivery (Dolan and Humphrey, 2001). In autos, advanced country suppliers such as Magna, Bosch, Lear, and Yazaki have set up global operations to support the network of final assembly plants that automakers have established to serve local markets (Sturgeon and Florida, 1999; Sutton, 2005). In electronics, lead firms such as Alcatel, Nortel, and Hewlett Packard have outsourced production to a set of huge, globally operating contract manufacturers including Celestica, Flextronics, Hon-hai, and Solectron (Sturgeon, 2002). In consumer goods and apparel, foreign companies do not merely buy what China produces and then resell it to North American consumers —Wal-Mart alone imported $15 billion worth of goods to North America from China in 2003—they actively shape the industrial transformation that has made the rise of China possible (Gereffi, 1994; Feenstra and Hamilton, 2006). The vast majority of exporting factories in Mainland China are run by firms from other economies such as Taiwan, Korea, Australia, Europe, Japan, and the United States; and most make products according to the detailed specifications set by non-Chinese firms such as Wal-Mart, Costco, Dell, and Nike. Clearly we need to look beyond trade and investment statistics to find out where the power in these global-scale production arrangements lie, and how these arrangements are changing.

Julia Lane of the National Science Foundation has likened the current state of qualitative industry research to the study of the natural world in the 16th and 17th centuries.[3] Curious researchers made detailed notes and drawings of what they could see of the vastness and variety around them, but there were few mechanisms for compiling the findings of individual researchers into larger pools of knowledge that could reveal broad patterns. Comparison of results came haphazardly with personal

---

[3]   These remarks were made at the MIT Working Group on Services Offshoring Workshop, held in Cambridge, Massachusetts on October 28, 2005.

48

communication between scholars and in the few forums, such as the British Royal Society, where researchers could present and debate their results. In this way classification systems gradually came into being and some of the mechanisms at work in nature were revealed.

Similarly, industry researchers have now had several decades to present, publish, and debate their research results, and more effort is now shifting to the construction of classification systems and to a search for the mechanisms that work to create the variety observed in the field.

The findings show that global integration is expressed differently in different industries and places. The precise patterns and effects of global integration, therefore, depend in large part on the technical and business characteristics that prevail in specific industries, and upon social and institutional characteristics of the places in which the nodes of global industries are embedded. For example, some industries, or parts of industries, are easier to fragment and globalize than are others.

What is needed now is a generic theory to explain the different patterns and to predict the outcomes associated with them.

From Global Commodity Chains to Global Value Chains

A pioneering step toward the development of an industry-independent, firm-level theory of global governance was taken in a chapter that Gary Gereffi wrote for the 1994 book he edited with Miguel Korzeniewicz, *Commodity Chains and Global Capitalism*, entitled 'The Organization of Buyer-driven Global Commodity Chains: How U.S. Retailers Shape Overseas Production Networks.' Building on the work of Hopkins and Wallerstein (1977, 1986), who highlighted the power of the state in shaping global production systems, or 'global commodity chains' (GCC's), exercised in large part in the form of tariffs and local content rules at the point where goods and investment crossed borders, Gereffi broadened the focus of the GCC framework to include the strategies and actions of firms. Gereffi's framework laid out four key structures that characterize

and shape GCCs (input-output, geographic, governance, and institutional) but one, governance, received the most attention, both from Gereffi and his immediate co-authors and from the many others that have made use of the GCC framework.[4]

One reason for the shift of focus to firm-level governance was the restricted ability of states to set tariffs and local content rules because of trade liberalization. However, Gereffi was one of the first scholars to argue convincingly that trade openness alone cannot explain the creation of industrial capabilities in developing countries. The best example of this is export-oriented industrialization in East Asia. It is commonly observed that the rapid growth of exports from developing countries has come with increased trade openness in the West. This is indeed a necessary condition for the export-oriented development that has been characteristic of the world's fastest growing economies, such as Taiwan, Korea, and China. But this development path has also been influenced and enabled by the competitive strategies of American, European, and Japanese firms, which established local operations, identified local firms as suppliers, transferred skills and technologies to them, invested in them, sold advanced equipment and services to them, and consistently ordered from them in a manner that pressured them to increase their industrial, technological, and logistical capabilities. As a

---

[4] The first two structures mentioned by Gereffi, input-output and geographic, are largely descriptive. Firms, and the larger networks and industries to which they contribute, can be described as an amalgam of value-added activities. Through simple, if painstaking, observational research, each of these activities can be located, both organizationally and geographically, yielding the first two of Gereffi's four structures. Governance and institutional structures are causal of the first two, and so require theoretical explanation. Institutional influences on the organizational and geographic structures in global chains remain to be adequately theorized, though literature from the field of political science on varieties within capitalism provide some guidance at the level of national-level institutions (e.g., Hall and Soskice, 2001). But clearly, supra-national institutions, such as the trade rules set in the context of the World Trade Organization, can have strong effects on the geographic and organizational patterns observed in global chains.

result, developing countries, especially in East Asia, were able to establish and upgrade a critical set of domestic technological and industrial capabilities with great rapidity. Liberalization has enabled the growth of international trade, but without the push from advanced economy firms seeking to tap capabilities in developing countries, the cross-border flows of goods and services would surely be more modest, in terms of both total volume and technological content, than they are today. Because firms from advanced economies have done so much to create capabilities in developing countries, they continue to control and guide many of the key industrial resources in the global economy, even those that they do not own.

Clearly, some firms exercise a greater degree control over the shape and extent of global production networks than others. The shift in focus was from the state to the actors in the chain, and their interrelationships, especially the relative power that "lead firms," the firms that place orders in global production networks, are able to exert to influence the actions of their affiliates and trading partners. Specifically, the GCC framework as adapted by Gereffi developed a key distinction between global chains that are "driven" by two kinds of lead firms: buyers and producers. The GCC framework usefully focused attention on the powerful role that large retailers, such as Wal-Mart, and highly successful branded merchandisers, such as Nike, have come to play in the governance of global production and distribution. Although "global buyers" typically own few, if any, of their own factories, the volume of their purchasing provides them with a huge amount of clout among their suppliers, power they have wielded to specify in great detail what, how, when, where, and by whom the goods they sell are produced. Extreme market power has also allowed global buyers to extract price concessions from their main suppliers. Supplier firms have responded by locating more of their factories in low-cost locations and working hard to extract price concessions from their own, upstream suppliers.

The GCC framework contrasted such "buyer-driven" chains with "producer-driven" chains, dominated by large manufacturing firms such as General Motors and IBM. Put sim-

ply, producer-driven chains have more linkages between affiliates of multinational firms, while buyer-driven chains have more linkages between legally independent firms. Underlying this distinction is the notion that buyer-driven chains turned out relatively simple products, such as apparel, house wares, and toys. Because innovation lies more in product design and marketing rather than in manufacturing know-how, it was relatively easy for lead firms to outsource production. In the more technology- and capital-intensive items made in producer-driven chains, such as autos and complex electronics, technology and production expertise were core competencies that needed to be developed and deployed in-house, or in tightly affiliated "captive" suppliers that could be blocked from sharing them with competitors.

Since Gereffi's seminal work was published in 1994, transnational giants have changed quite dramatically, outsourcing many activities and developing strategic alliances with competitors. In short, they have become less vertically integrated and more network-oriented. Better global standards in the realms of business processes and product characteristics, and the heavy application of information technology in areas such as design, manufacturing, service provision, supply-chain coordination, and materials management, have enabled increased outsourcing in producer-driven chains and made it possible, and more compelling, for firms to use these systems to streamline the linkages between buyers and suppliers in both producer- and buyer-driven chains. The result has been broad and rapid shifts in chain governance, where producers have become more buyer-like through outsourcing, and where the capabilities required to serve global buyers have been escalating rapidly. Today, global-scale networks of legally independent firms no longer make only simple items, but technology- and capital-intensive goods and services as well.

Because of these changes, there was a need to move beyond the GCC framework. Field research in a range of global industries revealed convergence in global value chain structure toward external networks, but not all of these industries were labour intensive, and this demanded more network types than buyer-

driven. Specifically, there are four new features in the governance of global-scale production networks that stimulated us to reconceptualize the key variables in global chain governance:

- Improvements in information technology and industry-level standards were enabling the codification of complex information, which in turn was easing the way for network forms of organization in technology-intensive industries.

- Flexible capital equipment was enabling the pooling of technology- and capital-intensive production in the same way that labour-intensive production could be pooled, again easing the way for network forms of organization in technology-intensive industries.

- Sophisticated supply-chain management tools were pushing labour-intensive industries up the technology curve.

- Increased outsourcing by manufacturing firms, and increased involvement in product definition by retailers (private label) were blurring the distinction between buyers and producers.

So work began to develop a new theory for understanding, explaining, and predicting firm-level governance patterns in the global economy. The result was the Global Value Chains (GVC) framework, developed by a network of scholars, practitioners, policy-makers, and NGO activists over a period of several years.[5] The central questions we asked were: How are specific industries coordinated at a global scale? What are the key variables that contribute to these governance patterns? What patterns of global value chain governance can be expected when these variables change? It is our attempt to answer this last question that set the GVC framework apart. Instead of a static typology, we sought to develop an operational conceptual model. In other words, changing the value of the variables should yield distinct and predictable patterns of global value chain governance.

See http://www.globalvaluechains.org for a summary of the GVC Initiative as well as a list of related researchers and publications.

To sum up, the GCC framework was extremely valuable because it shed light on the new and powerful role that retailers and branded merchandisers were playing in global production networks. But the producer/buyer-driven typology was static. It provided no mechanisms to account for shifts in barriers to network entry brought on by technological change or firm- and industry-level learning. As empirical changes forced us to take a more dynamic view of the governance patterns in global production network, two things became clear:

1) there was a convergence of chain governance away from the producer-driven variant toward external, non-equity networks, and

2) the buyer-driven type could not characterize all of the network types being observed in the field.

The Global Value Chains Framework

The GVC framework was first published in an article entitled 'The Governance of Global Value Chains' in the journal *Review of International Political Economy* (Gereffi, Humphrey, and Sturgeon, 2005). The article sought to both account for the recent observed changes in the organization of the global economy and to build a more theoretically-grounded approach to explaining and predicting firm-level governance patterns in geographically separate economic activities. As such, we were consciously extending and refining the GCC framework.

The GVC framework is a tool kit for understanding how activities are linked across great distances in the global economy. The main object of inquiry is the nature and content of the link between value-added activities. For simplicity's sake, we began with two kinds of firms, lead firms, or order makers, and suppliers, or order takers. Much of the literature that seeks to create governance categories by examining the linkages between buyers and sellers in the global (or local) economy identify only two options: market or hierarchy (Williamson, 1975). Firms either invest offshore directly or buy goods and services from foreign firms. A smaller body of literature has noted the prevalence of network forms of organization where there is

some form of "explicit coordination" beyond simple market transactions but which fall short of vertical integration (Powell, 1990; Adler, 2001). While the insights from this "network" literature are useful, our field research convinced us that not all networks are the same. We identified four kinds of transactional linkages between lead firms and suppliers, market, modular, relational, and captive, and summarized all manner of intra-firm linkages as hierarchical. This yields five types of linkages, which, assuming that all linkages in a given chain of activities are governed similarly, aggregate into five ideal types of GVC governance.[6] In essence, the GVC framework specifies three types of network governance (modular, relational, and captive) along with the two traditional modes of economic governance (markets and hierarchies). The characteristics of the five GVC governance types are summarized in Table 2.

---

[6] Obviously, in the real world, a given value chain will display a mix of governance forms. To complicate matters further, in-house linkages can also take a variety of forms.

Table 2. Five Forms of Global Value Chain Governance

1. *Markets*. Markets are the simplest form of GVC governance. GVCs governed by markets contain firms and individuals with little interaction beyond exchanging goods and services for money. The central governance mechanism is price. The linkages between value chain activities are not very "thick" because the information that needs to be exchanged and knowledge that needs to be shared are relatively straightforward.

2. *Modular value chains*. This is the most market-like of the three network-style GVC governance patterns. Typically, suppliers in modular value chains make products or provide services to a customer's specifications. Suppliers in modular value chains tend to take full responsibility for process technology and often use generic machinery that spreads investments across a wide customer base. This keeps switching costs low and limits transaction-specific investments, even though buyer-supplier interactions can be very complex. Linkages are necessarily thicker than in simple markets because of the high volume of information flowing across the inter-firm link, but at the same time codification schemes and the internalization of coherent realms of knowledge in value chain "modules," such as design or production, can keep interactions between value chain partners from becoming highly dense and idiosyncratic.

3. *Relational value chains*. In this network-style GVC governance pattern we see mutual dependence regulated through reputation, social and spatial proximity, family and ethnic ties, and the like. The most obvious examples of such networks are in specific communities, or "industrial districts," but trust and reputational effects can operate in spatially dispersed networks as well. Since trust and mutual dependence in relational GVCs take a long time to build up, and since the effects of spatial and social proximity are, by definition, limited to a relatively small set of co-located firms, the costs of switching to new partners tends to be high. Dense interactions and knowledge sharing are supported by the deep understanding value chain partners have of one another, but unlike the codification schemes that enable modular networks, these "short-cuts" tend to be idiosyncratic and thus difficult and time-consuming to re-establish with new value chain partners.

4. *Captive value chains*. In this network-style GVC governance pattern, small suppliers tend to be dependent on larger, dominant buyers. Depending on a dominant lead firm raises switching costs for suppliers, which are "captive." Such networks are frequently characterized by a high degree of monitoring and control by the lead firm. The asymmetric power relationships in captive networks force suppliers to link to their customer in ways that are specified by, and often specific to a particular customer, leading to thick, idiosyncratic linkages and high switching costs all round.

5. *Hierarchy*. This governance pattern is characterized by vertical integration (i.e., "transactions" take place inside a single firm). The dominant form of governance is managerial control.

6. The *exclusion of suppliers* from the chain. If there is low complexity and a high possibility for codification, and suppliers still do not have the capabilities to meet the requirements of buyers, then it is likely that they will be excluded from the chain. While this does not generate a global value chain type, *per se*, it is a situation that is quite common, and with requirements for suppliers increasing, perhaps increasingly likely to occur (Sturgeon and Lester, 2004).[7]

When would we expect each of these five governance forms to occur? From our field research, reading, and discussions, we have identified three key variables:
1) the *complexity* of the information exchange required to complete the transaction;
2) the degree to which the information can be expressed formally, or its *codifiability*; and
3) the level of competence in the supplier relative to the transaction.

The three variables are summarized in Table 3.

---

[7]   It is the exclusion of developing country suppliers that has motivated us, more than any other factor, to construct this theory of global value chains. If the framework can help to make sense of value chain governance patterns, then it can be used as a basis for the development of more effective policies for industrial upgrading, especially those aimed at rectifying situations of supplier exclusion.

Table 3. Three Key Variables in Global Value Chain Governance

1. *The complexity of transactions.* More complex transactions require greater interaction among actors in GVCs and thus stronger forms of governance than simple price-based markets. Thus, complex transactions will likely to be associated with one of the three network governance patterns (modular, relational, or captive) or integrated within a single firm (hierarchy).

2. *The codifiability of transactions.* In some industries schemes have been worked out to codify complex information in a manner in which data can be handed off between GVC partners with relative ease, often using advanced information technologies. If suppliers have the competence to receive and act upon such codified information, and if the codification schemes are widely known and widely used, then we would expect to see modular value chains emerge. If not, then lead firms might either keep the function in-house, leading to more vertical integration (hierarchy) or outsource it to a supplier that they tightly control and monitor (the captive network type) or have a dense, idiosyncratic relationship with suppliers (the relational governance type).

3. *The competence of suppliers.* The ability to receive and act upon complex information or instructions from lead firms requires a high degree of competence on the part of suppliers. Only then can the transfer of complex but codified information be achieved (as in modular networks) or intense interaction be worthwhile (as in relational networks). Where competent suppliers do not exist, lead firms either must internalize the function (hierarchy) or outsource it to suppliers that they tightly monitor and control (captive suppliers).

Furthermore, if one of these three variables changes, then value chain governance patterns tend to change in predictable ways. For example, if a new technology renders an established codification scheme obsolete, we would expect, all other thing being equal, modular value chains to become more relational, and if competent suppliers cannot be found, then captive networks and even vertical integration would become more prevalent. Conversely, rising supplier competence would tend to push captive governance more toward the relational type and better codification schemes might prepare the ground for modular governance. The five global value chain governance types, along with the values of the three variables that determine them, are shown in Figure 1. The five types of global value chain governance are derived from ascribing binary (high or low) values to the three key variables: 1) complexity of inter-firm transactions; 2) the degree to which this complexity can be mitigated

through codification; and 3) the extent to which suppliers have the necessary capabilities to meet the buyers' requirements. Each governance type provides a different trade-off between the benefits and risks of outsourcing. As shown in the last column of Figure 1, the governance types comprise a spectrum running from low levels of explicit coordination and power asymmetry between buyers and suppliers, in the case of markets, to high levels of explicit coordination and power asymmetry between buyers and suppliers, in the case of hierarchy.

Figure 1. The Global Value Chains Framework

	Key Variable			
Govern-ance Type	*Complexity of transac-tions*	*Ability to codify transactions*	*Capabili-ties in the supply base*	*Degree of explicit coordination and power asymmetry*
Market	Low	High	High	
				Low ↑
Modular	High	High	High	
Relational	High	Low	High	
Captive	High	High	Low	
				High ↓
Hierarchy	High	Low	Low	

Note: There are eight possible combinations of the three variables. Five of them generate global value chain types. The combination of low complexity of transactions and low ability to codify is unlikely to occur. This excludes two combinations. If the complexity of the transaction is low and the ability to codify is high, then low supplier capability would lead to exclusion from the value chain. While this is an important outcome, it does not generate a governance type *per se*.

The theory of GVC governance presented here combines key insights from a variety of disciplines. First, from institutional economics we accept the importance of asset specificity as a key problem in inter-firm relationships (Williamson 1985). Firms that have assets that are specific to a single trading partner create risks of hold-up that must somehow be solved. Institutional economics offers vertical integration as the solution, yielding the markets and hierarchies dichotomy. While we agree that vertical integration is one way to solve the problem of asset

specificity, literatures from several other fields convince us that there are other solutions. From the economic, sociology and economic geography literatures we learn that trust and reputation, built up over time through repeat transactions and enabled by social and spatial proximity, can offset the risks of asset specificity without vertical integration (Piore and Sabel, 1984; Granovetter, 1985; Storper, 1995). The resource view of the firm prevalent in the management literature (Penrose, 1959) convinces us that some business functions can be very difficult to acquire or to develop in-house, and so firms must continue to source them externally even if asset specificity is present. Finally, from concepts such as "open innovation," "platform management," and "modularity" developed in recent literature from the field of business history we learn that firms can engineer their way out of problems of asset specificity by standardizing and codifying information at specific "pinch points" in the chain of value added activities (Baldwin and Clark, 2000; Chesbrough and Kusunoki, 2001; Gawer and Cusumano, 2002).

The GVC framework is not intended to provide a grand theory of economic development, or even a full theory of the forces that shape global integration, but a transaction-, firm- and industry-centric theory of governance among the firm- and establishment-level actors in the chain. As such it cannot provide a full accounting of the governance characteristics of global value chains. It can, however, provide a bottom-up, research-driven method for accounting for observed global value chain governance characteristics as well as those that are predicted to arise *absent other factors and influences*. In this way, the GVC framework can provide researchers and policy-makers with a useful and relatively simple first-cut: if the value chain governance patterns that are predicted by the theory are not observed empirically, this provides a strong indication that forces external to the chain, such as national institutions or international trade rules, are playing a large role. In this way, the framework can provide insight into, but not a full accounting for, the observed features of global integration. While implications for policy are numerous, they depend wholly on context, and I will not attempt to fully elaborate them here. To provide just one

example, relational value chains, or more accurately the relational segments of value chains, tend to be less geographically mobile, not only because they require frequent interaction to develop and exchange tacit knowledge, but because they are often governed by the shared expectations of trust and reputation developed over long periods of time.

Modularity in Global Value Chains

Of all the governance forms generated by the global value chains framework, modular value chains are the most novel because they are enabled by very recent advancements in information technology, robust industry standards, and global-scale data communications systems. They also warrant our attention because they are extremely fluid organizationally and geographically and so readily allow the production of complex goods and services to be fragmented and geographically dispersed. Finally, modular value chains, because they are based on the exchange of information according to well known standards, are very fluid relative to other forms of GVC governance that require more time to establish and re-establish efficient inter-firm linkages. This fluidity, in terms of where specific value chain activities are carried out and by whom, has contributed to the acceleration of global integration mentioned earlier. Because the rise of value chain modularity has obvious implications for policy, it is worth examining its features and precursors in some detail.

Value chain modularity is based on functional specialization, formalization of value chain linkages, and an increase in the scale and geographic reach of each horizontal segment —or 'module'—of the value chain (Baldwin and Clark, 2000; Sturgeon, 2002; Takaishi and Fujimoto, 2001; Langlois, 2003; Prencipe et al., 2003; Gereffi et al., 2005; Sturgeon and Lee, 2005). In modular value chains distinct breaks in the sequence of activities tend to form at points where information regarding product and process specifications can be highly formalized. Activities tend to remain tightly integrated and based on tacit linkages within functionally specialized nodes of "relational" activity. Within these relational nodes tacit knowledge is cre-

ated, exchanged, and processed by establishments and workers who tend to be co-located. Between these nodes, however, linkages are eased by the application of widely agreed-upon protocols and standards. Codified linkages allow value chain modules to more easily be located at great distance.

Codification and standardization have helped to create simplified and reliable methods for transmitting detailed product and production specifications along the value chain, and for keeping track of large, complex projects with participants in diverse locations and organizations. As Coase (1937) perceived, this sort of formalization of information at the inter-firm link lowers transaction costs, allowing firms to more easily purchase inputs on the market. However, the content of this information exchange, as Coase envisioned it, consisted largely of price data. The high volume of non-price data flowing across the inter-firm link differentiates modular value chains from simple markets. Because of this complexity it is not unusual that additional engineering and coordination is required to complete a transaction. The hand-off of product and process specifications between firms need not be perfectly clean, but only relatively so for modular value chains to function.

Specifically, the key business processes that have been formalized, codified, standardized, and computerized are product design (e.g., computer aided design), production planning and inventory and logistic control (e.g., enterprise resource planning), as well as various aspects of the production process itself (e.g., assembly, test and inspection, material handling). Furthermore, because it is "platform independent," the Internet has provided an ideal vehicle for sharing and monitoring the data generated and used by these systems. Such technologies and practices are at the core of value chain modularity. It is the formalization of information and knowledge at the inter-firm link, and the relative independence of the participating firms that gives value chain modularity its essential character: flexibility, resilience, speed, and economies of scale that accrue at the level of the industry rather than the firm (Sturgeon and Lee, 2005). Modular linkages between relational nodes of tacit activity can exist within a single firm, but only when activities are

outsourced can scale economies build up beyond the level of the firm (Langlois and Robertson, 1995).

Value chain modularity introduces risks as well as benefits for participating firms. Responsiveness may suffer as contracts are hammered out. There is potential for intellectual property and other sensitive information about product features, pricing, production forecasts, and customers to leak to competitors through shared suppliers. The ability of lead firms to innovate and design successive product generations may suffer from the atrophying of manufacturing and component knowledge, a problem that has been referred to by Chesbrough and Kusunoki (2001) as the 'modularity trap.' Reliance on standard interfaces may lead to the use of standard components, leading in turn to a loss of product distinctiveness. Shared and overlapping inventory resident in supplier organizations can lead to distortions and tracking problems that introduce waste. One unavoidable issue is that independent firms in buyer-supplier relationships often have competing interests.

In the American electronics industry, value chain modularity took shape during the late 1980s and early 1990s. Because many established firms had in-house manufacturing and components divisions, this change required the break-up of vertically integrated corporate structures and the aggregation of cast off activities in suppliers. Hewlett Packard and IBM led the way, selling most of their worldwide manufacturing infrastructure to contract manufacturers such as Solectron and Flextronics, or spinning off internal divisions as merchant contract manufacturers, as IBM did with its Toronto manufacturing complex in 1997, creating the contract manufacturer Celestica. Another source of growth in contract manufacturing was increased business from newer firms that never built up internal manufacturing divisions, such as the Internet switch company Cisco and the computer workstation and server firms Sun Microsystems and Silicon Graphics (Sturgeon, 2002).

Outsourced circuit board and final assembly of commercial electronics (products for the medical, automotive, communications, military, corporate computing markets) was mostly transferred to contract manufacturers based in North America, spe-

cifically the big five 'electronics manufacturing services' (EMS) firms Flextronics, Solectron, Sanmina-SCI, Jabil, and Celestica, while the assembly and even some of the design of notebook and desktop personal computers were outsourced to 'original equipment' and 'original design' (OEM and ODM) contract manufacturers based in Taiwan, such as Quanta, Compal, Inventec, Hon Hai (Foxconn), and Wistron, the contract design and manufacturing arm of Acer. By the end of the 1990s, much of the manufacturing capacity of the Taiwan-based contract manufacturers had shifted to mainland China, and the big five United States-based contract manufacturers had established a global-scale network of factories (Sturgeon and Lester, 2004). At the level of components, the 1990s was a time of rapid growth among 'fabless' semiconductor design firms and the semiconductor foundries (chip manufacturing plants) that served them, such as the Taiwan-based TSMC and UMC, as well as IBM (Linden and Samaya, 2003). By the end of the 1990s, modular value chains in the electronics industry were highly developed and global in scope.

An important aspect of global integration, then, is the globalization of the supply-base. As more work has been handed off to independent suppliers and service providers that are tied to lead firms though modular value chain linkages, two things have happened. First, the largest suppliers have achieved unprecedented economies of scale and scope by pooling resources across a broad customer base. As a result, some have increased the geographic scope of their business to the point where they have global-scale operations. So the multinational firm continues to be a powerful force regardless of the fragmentation and re-bundling in the value chain. Second, the existence of highly competent independent local and global suppliers has lowered the barriers to globalization for firms, including small and medium-sized firms, which have not yet shifted any activities offshore. When smaller, less competent firms begin to look outside of their own companies and perhaps even offshore for key inputs, markets are created for a broader array of functions to be outsourced, and this drives suppliers to bundle additional functions and further increase their range of competencies.

Conclusions

Integrative trade is being driven by vertical fragmentation, where firms specialize in providing specific bundles of goods and services to a larger network of firms. This bundling and packaging of functions by suppliers lowers barriers to global sourcing yet further, setting in motion a cycle of increased supply-base competence and increased outsourcing and offshoring of which we clearly have not seen the end. The "co-evolutionary" view of global-scale economic integration (Sturgeon and Lee, 2005) that is embedded in the GVC framework emphasizes that the patterns of globalization that have already developed work to alter future patterns. For example, we must consider the possibility that the pace of globalization observed in manufacturing industries since the 1970s will be a poor indicator of what is likely to happen in services. The offshoring of computer hardware production began at a time when the firms in societies receiving this new business had few capabilities. International communications systems were slow, unreliable, of limited functionality, and very costly to use. Services offshoring, by contrast, is expanding with the infrastructure, firm capabilities, and business models that have been established, tested, and refined in support of global manufacturing already in place. Integrative trade in services, then, will flow down the well-trodden avenues in the global economy that were put in place largely to support global-scale goods production: across highly functional and low-cost broadband communications systems, through cross-border business relationships that have now been in place for decades, according to business models regarding outsourcing and offshoring that have been worked out in exquisite detail, and through firms with huge, well established multinational operations. Looking to the future, we cannot and should not pretend to know precisely how much or what kind of economic activity will flow across these pathways, but we cannot afford to be complacent. The long-term prospects for any country may be less certain given the vastly altered playing field on which global integration is unfolding.

The lessons for research and policy are numerous. New thinking is needed to develop useful insights into the character and implications of our increasingly globally integrated national economies. Long cherished notions and responses may need to be set aside, to be replaced or at least supplemented with new theoretical frameworks and policy initiatives. Perhaps the most pressing need is for new kinds of data to be collected, data that shed light on the important questions of power and coordination within global value chains.

References

Adler, Paul. 2001. "Market, Hierarchy, and Trust: The Knowledge Economy and the Future of Capitalism." *Organization Science* 12(2) (March-April), pp. 215-234.

Baldwin, Carliss, and Kim Clark. 2000. *Design Rules*, Cambridge, MA: MIT Press.

Bardhan, Ashok Deo; and Cynthia Kroll. 2003. "The New Wave of Outsourcing." Fisher Center for Real Estate & Urban Economics, Research Report #1103, November 2. Online at http://repositories.cdlib.org/iber/fcreue/reports/1103.

Batt, Rosemary; Virginia Doellgast, and Hyunji Kwon. 2005. "A Comparison of Service Management and Employment Systems in US and Indian Call Centers." Paper prepared for the Brookings Trade Forum 2005: Offshoring White-Collar Work — The Issues and the Implications. Susan M. Collins and Lael Brainard, eds. November 10, Washington, DC. Online at http://www.brookings.edu.

Berger, Suzanne, and the MIT Industrial Performance Center. 2005. *How We Compete*. New York: Doubleday.

Bernard, Andrew B., J. Bradford Jensen and Peter K. Schott. 2005a. "Survival of the Best Fit: Exposure to Low-Wage Countries and the (Uneven) Growth of U.S. Manufacturing Plants." *Journal of International Economics*.

Bernard, Andrew B., J. Bradford Jensen and Peter K. Schott. 2005b. "Importers, Exporters and Multinationals: A Portrait of Firms in the U.S. that Trade Goods." NBER Working Paper 11404.

Bhagwati, Jagdish. 2004. "The Muddles over Outsourcing." *Journal of Economic Perspectives* (Fall).

Blinder, Alan S. 2005. "Fear of Offshoring." CEPS Working Paper No. 119, Princeton University, December.

Breznitz, Dan. Forthcoming 2007. *Innovation and The State*. Yale University Press.

Chesbrough, H. and K. Kusunoki. 2001. "The Modularity Trap: Innovation, Technology Phase Shifts, and the Resulting Limits of Virtual Organizations." In L. Nonaka and D. Teece,

eds., *Managing Industrial Knowledge*. London: Sage Press, pp. 202-230.

Coase, R. 1937. "The nature of the firm." Economica. 4: 386-405.

Dicken, P. 1992. *Global Shift: The Internationalisation of Economic Activity.* London: Paul Chapman Publishing.

Dolan, C., and J. Humphrey. 2000. "Governance and trade in fresh vegetables: the impact of UK supermarkets on the African horticulture industry." *Journal of Development Studies* 37(2): 147-176.

Dolan, C., and J. Humphrey. 2001. "Changing governance patterns in the trade in fresh vegetables between Africa and the United Kingdom." Mimeo. Brighton: Institute of Development Studies.

Dossani, Rafiq, and Martin Kenney. 2003. "Lift and Shift; Moving the back office to India." Work in Progress (Sept.). *Information Technologies and International Development*, 1(2) (Winter), pp. 21–37. Massachusetts Institute of Technology, Cambridge, MA.

Dossani, Rafiq, and Martin Kenney. 2004. "The Next Wave of Globalization? Exploring the Relocation of Service Provision to India." Berkeley Roundtable on the International Economy Working Paper 156, September. University of California Berkeley, Berkeley, CA.

Dossani, Rafiq, and Martin Kenney. 2005. "Moving Services Offshore: A Case Study of a U.S. High-Technology Firm." Report was prepared for the European Union - U.S. Department of Labor seminar *Offshoring of Services in ICT and Related Services*.

Feenstra, Robert and Gary Hamilton. 2006. *Emergent Economies, Divergent Paths: Economic Organization and International Trade in South Korea and Taiwan*. Structural Analysis in the Social Sciences series. Marc Granovetter, ed. Cambridge, UK: Cambridge University Press.

Feenstra, Robert. 1998. "Integration of Trade and Disintegration of Production in the Global Economy." *Journal of Economic Perspectives* 12(4), pp. 31–50.

Freeman, Richard. 2005. "China, India and the Doubling of the Global Labor Force: Who pays the price of globalization?" The *Globalist*, June 3.

Gawer, Annabelle and Michael Cusumano. 2002. *Platform Leadership; How Intel, Microsoft, and Cisco Drive Innovation.* Boston, Massachusetts: Harvard Business School Press.

Gereffi, G., J. Humphrey, and T. Sturgeon. 2005. The Governance of Global Value Chains, *Review of International Political Economy* 12(1) (February), pp. 78-104.

Gereffi, Gary. 1994. "The organization of buyer-driven global commodity chains: how U.S. retailers shape overseas production networks." In G. Gereffi and M. Korzeniewicz, eds., *Commodity Chains and Global Capitalism.* Westport: Praeger, pp. 95-122.

Gomory, R.E., and W.J. Baumol. 2001. *Global Trade and Conflicting National Interests.* Cambridge, Mass.: MIT.

Granovetter, Mark. 1985. "Economic Action and Social Structure: the Problem of Embeddedness." *American Journal of Sociology* 91: 481-510.

Grimes, Anne. 2004. "Venture Firms Seek Startups that Outsource." *Wall Street Journal.* Section B1, Column 5, April 2.

Hall, Peter and David Soskice, eds. 2001. *Varieties of Capitalism.* Oxford, UK: Oxford University Press.

Harrison, Ann E. and Margaret S. McMillan. 2006. "Outsourcing Jobs? Multinationals and U.S. Employment.." NBER Working Paper 12372. National Bureau of Economic Research. Cambridge, MA. July. Online at http://www.nber.org/papers/w12372.

Hopkins, T. and I. Wallerstein. 1977. "Patterns of Development of the Modern World-System." *Review* 1(2), 11-145.

Hopkins, T. and I. Wallerstein. 1986. "Commodity chains in the world economy prior to 1800." *Review* 10(1), 157-170.

Hummels, D., D. Rapoport, and K.-M Yi. 1998. "Vertical specialization and the changing nature of world trade." In Federal Reserve Bank of New York, *Economic Policy Review* 4(2), 79-99.

Jarmin and Miranda. 2002. "The Longitudinal Business Database." Mimeo. Census Bureau. July 16. Accessed at

www.ces.census.gov/paper.php?paper=101647, October 15, 2002.

Kletzer, Lori. 2002. *Imports, Exports, and Jobs: What Does Trade Mean for Employment and Job Loss?* Kalamazoo, Michigan: W.E. Upjohn Institute for Employment Research.

Langlois, R. and P. Robertson. 1995. *Firms, Markets and Economic Change.* London: Routledge.

Langlois, Richard. 2003. "The Vanishing Hand: The Changing Dynamics of Industrial Capitalism." *Industrial and Corporate Change* 12(2), 351-385.

Levy, Frank, and Ari Goelman. 2005. "Offshoring and Radiology." In Susan M. Collins and Lael Brainard, eds., *Offshoring White-Collar Work - Issues and Implications.* Brookings Trade Forum, Washington, DC.

Levy, Frank, and Richard J. Murnane. 2004. *The New Division of Labor: How Computers Are Creating the Next Job Market.* Princeton, NJ: Princeton University Press.

Linden, Greg and Deepak Somaya. 2003. "System-on-a-Chip Integration in the Semiconductor Industry: Industry Structure and Firm Strategies." *Industrial and Corporate Change* 12(3), 545-557.

Mieszcowski, Katharine. 2003. "Moving to India Is Not a Luxury. It Is a Necessity." *Salon.Com*, December 17, 2003.

Penrose, Edith. 1959. *The Theory of the Growth of the Firm.* Oxford: Basil Blackwell.

Piore, M. and C. Sabel. 1984. *The Second Industrial Divide.* New York: Basic Books.

Powell, W. 1990. "Neither Market Nor Hierarchy: Network Forms of Organization." *Research in Organizational Behavior* 12, 295-336.

Prencipe, Andrea, Andrew Davies, and Mike Hobday, eds. 2003. *The Business of Systems Integration.* Oxford: Oxford University Press.

Samuelson, Paul A. 2004. "Where Ricardo and Mill rebut and confirm arguments of mainstream economists against globalization." *Journal of Economic Perspectives* 18(3).

Storper, Michael. 1995. "The Resurgence of Regional Economies, Ten Years Later: The Region as a Nexus of Untraded

Dependencies." *European Urban and Regional Studies* 2(3), 191-221.

Sturgeon, T., and R. Florida. 2004. "Globalization, deverticalization, and employment in the motor vehicle industry." In M. Kenney and B. Kogut, eds., *Creating Global Advantage.* Stanford, CA: Stanford University Press.

Sturgeon, Timothy and Ji-Ren Lee. 2005. "Industry Co-Evolution: Electronics Contract Manufacturing in North American and Taiwan." In S. Berger and R. Lester, eds., *Global Taiwan: Building Competitive Strengths in a New International Economy.* New York, London: M.E. Sharpe, pp. 33-75.

Sturgeon, Timothy and Richard Lester. 2004. "The New Global Supply-base: New Challenges for Local Suppliers in East Asia." Chapter two in *Global Production Networking and Technological Change in East Asia.* Shahid Yusuf, Anjum Altaf and Kaoru Nabeshima. eds. Oxford: Oxford University Press.

Sturgeon, Timothy. 2002. "Modular Production Networks. A New American Model of Industrial Organization." *Industrial and Corporate Change* 11(3), 451-496.

Sturgeon, Timothy; Frank Levy, Clair Brown, J. Bradford Jensen; and David Weil. 2006. *MIT Industrial Performance Center Services Offshoring Working Group Final Report.* Industrial Performance Center Working Paper. Online at http://web.mit.edu/ipc/www/IPC_Offshoring_Report.pdf.

Sutton, John. 2006. "The Globalization Process: Auto-Component Supply Chains in China and India." In François Bourguignon, Boris Pleskovic, and André Sapir, eds., *Are We on Track to Achieve the Millennium Development Goals?* Annual World Bank Conference on Development Economics –Europe 2005. World Bank and Oxford University Press, pp. 185-213.

Swenson, Deborah. 2005. "Overseas assembly and country sourcing choices." *Journal of International Economics* 66(1), 107-130.

Takeishi, Akira and Takahiro Fujimoto. 2001. "Modularization in the Auto Industry: Interlinked Multiple Hierarchies of

Product, Product, and Supplier Systems." *International Journal of Automotive Technology and Management* 1(4), 379-396.

Williamson, Oliver E. 1975. *Markets and Hierarchies: Analysis and Antitrust Implications.* New York: The Free Press.

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism.* New York, NY: Free Press.

Wilson, Ron. 2003. "Venture Capitalist Explains New Rules for IC Startups." *EETimes,* January 16.

Yeats, A. 2001. "Just how big is global production sharing?" In S. Arndt and H. Kierzkowski, eds., *Fragmentation: New Production Patterns in the World Economy,* Oxford: Oxford University Press, pp. 108-143.

Data Issues on Integrative Trade between Canada and the US Measurement Issues for Supply Chains

Art Ridgeway [*]

Introduction

Reduced barriers to the flow of capital, goods, and services, combined with rapid advances in communication and transportation technology, have for many years been fostering increased specialization of production activity, and this trend continues. The same factors have also led large firms, particularly multinational firms, to reorganize how they manage their operations. Firms are increasingly focusing on supply chain management and the choice between make or buy for intermediate inputs and, increasingly, service inputs. Within the make-buy decision are decisions on the location of supply, domestic or international.

These factors have also encouraged the development or emergence of a number of economies that are now growing rapidly by providing lower cost alternatives for many production activities. This includes the so-called BRIC countries (Brazil, Russia, India, and China) and the eastern European economies. For example, China is now the second largest source of Canadian merchandise imports and Brazil is now one of Canada's top ten partner countries in both inward and outward foreign direct investment.

[*] Director of the Balance of Payments Division of Statistics Canada. This paper was originally prepared for the conference *Integrative Trade between Canada and the United States — Policy Implications*, organized by the Centre For Trade Policy and Law, held at Ottawa University, December 6, 2006. The views expressed in this paper are those of the author and not those of Statistics Canada.

The organization of some multinational enterprises now exhibits a separation of the management of operations and the legal structure of the enterprise. Operations can be managed in units that cross multiple countries, while there are, necessarily, separate legal structures for each country. While this dual structure allows the enterprise to efficiently manage its operations and respond to regulatory and fiscal requirements, neither structure conforms to that required to produce economic data in support of policy. In addition, these structures increase the importance of intra-firm transactions where transfer pricing may be an issue.

A recent international study by a group of national statistical agencies that attempted to coordinate the collection of data for a sample of multinational enterprises revealed that these multinational enterprises are very sensitive about the confidentiality of their data. They were quite concerned with efforts to reconcile across countries how they reported to the participating statistical agencies.

This paper looks at some of the implications of the globalization trend for economic statistics particularly with respect to the fragmentation of the production process and international trade. The annex provides a brief description of a range of statistical programs and current initiatives at Statistics Canada that are aimed at shedding light on the ongoing trend of globalization of production around the world economy.

Implications of Globalization Trends for Economic Statistics

These ongoing changes have led to the demand for new measures of economic activity and have affected how some traditional data series are gathered and interpreted. The drivers of the growth in globalization have been innovation and changing business structures and practices. Data on research and development activities have been available for many years, but more recently there have been demands for a broader set of data on innovation and the commercialization of new knowledge, and on the use of new electronic and other business practices. Statis-

tics Canada's science, innovation, and electronic information program continues to evolve in response to these new demands. In addition, there have been recent demands to directly measure the globalizing structural changes that are occurring. Offshoring is a phenomenon that users are seeking information on but that is not easily measured. Diagram 1 below presents a simple illustration of the relationship between the global enterprise, domestic production, outsourcing and offshoring.

Diagram 1: Global enterprises, production, outsourcing and offshoring

Global Enterprise

| Domestic in-house production | International in-house sourcing |
| Domestic outsourcing | International outsourcing |

Domestic production

Offshoring: International production

Outsourcing: Independent suppliers
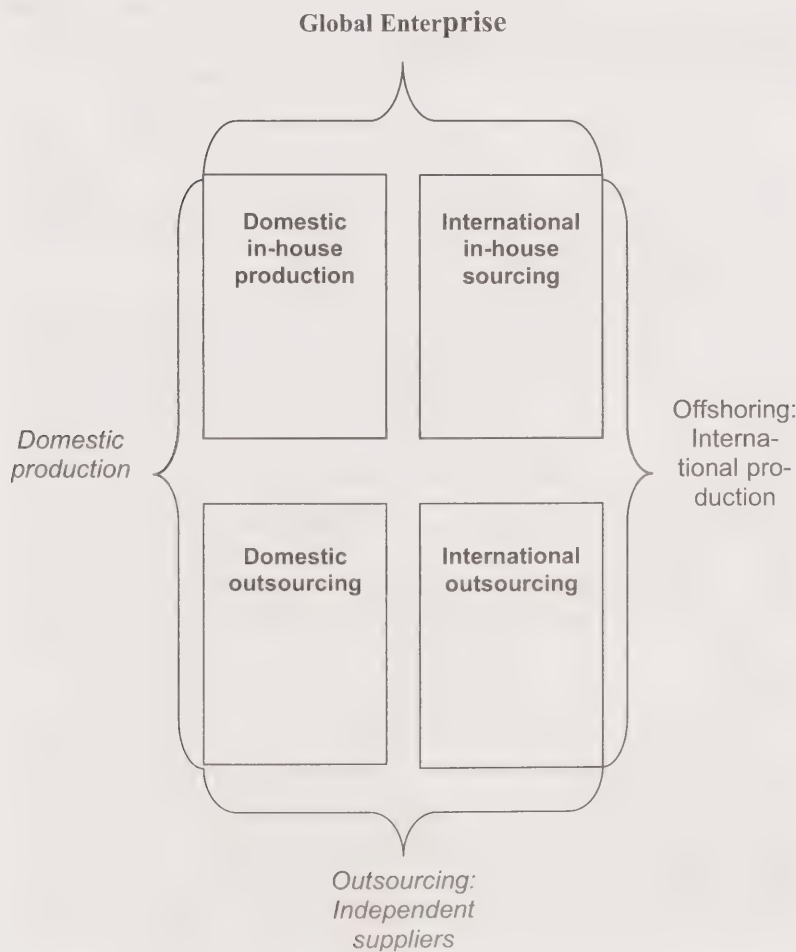
This diagram illustrates the broadest concepts of outsourcing and offshoring. Some analysts use these terms in a more restricted sense, defining them to refer just to those activities that have moved out of in-house or domestic production to independent or international production. Measuring outsourcing or offshoring events as defined by the narrow definition is very demanding, as this involves identifying decisions made at a specific point in time.

However, the most profound implications of increasing globalization may be in the challenges of measuring some of the core economic indicators. Moving productive activity out of the top left hand corner of the diagram to any of the three other quadrants implies that the ratio of gross flows to value added will increase. Factors that are increasing the measurement challenges include:

- the increasing value of service flows
- the increasing importance of flows internal to the global enterprise
- the increasing proportion of physical flows that are not coincident with changes in ownership

It used to be the case that when goods crossed borders they almost always changed ownership. This is no longer true. This separation of the ownership of goods and the cross-border flows in a multiple step production process is causing major valuation challenges. While there have been occasional instances of this for some time, the separation of ownership and production flows seems to be increasing. Collecting data from legal structures will generally reflect the ownership of resources whereas collection from operating units is more likely to reflect the operational flows of resources. While both are important for the full articulation of the economic data system, the reconciliation of data from the two separate structures is becoming more difficult.

These challenges and the growing demand for new policy relevant data come at a time of declining response rates for many business surveys. Fortunately, the increased use of administrative data for 'simple' businesses has greatly reduced the burden for the small and medium end of the population, while allowing an improved coverage of these units. However, there

are a few hundred of the largest enterprises in the country that account for a very large part of economic activity, and the deterioration in their rates of response to surveys is a concern. While declining response to surveys is not uniquely related to globalization per se, the fact that the non-respondents are often large globally oriented firms makes it difficult to get the data needed to measure globalization.

In addition, the growing importance of the emerging economies in Asia and South America pose challenges for bilateral comparisons as most of these countries have less well-developed statistical systems. Bilateral comparisons of data show large differences but, given the state of development of these statistical systems, it is difficult to assess the reasons for the discrepancies.

International Efforts

International organizations have been active in expanding the conceptual basis for measuring activity related to globalization. The OECD Technological Balance of Payments Manual has been in use since its release in 1990. More recently, the OECD has provided the OECD Handbook on Economic Globalisation Indicators and the related publication, Measuring Globalization: OECD Economic Globalisation Indicators.

Other statistical manuals currently under revision will provide improved links to the new measures of globalization. In particular, the SNA manual[1], the BOP manual[2], and the OECD Benchmark Definition of Foreign Direct Investment will have additional material on globalization. It has been announced recently that the manual on Statistics of International Trade in Services will also be updated to harmonize with the new material in the core manuals, in-

---

[1] For additional information on the updating of SNA 93 see United Nations National Accounts Section, Towards 1993 SNA Rev.1 http: unstats.un.org unsd nationalaccount snarev1.asp

[2] Additional information on the revision of the BOP Manual see Revision of the Fifth Edition of the IMF's Balance of Payments Manual http://www.imf.org/external/np/sta/bop/bopman5.htm

cluding that on globalization. Statistics Canada has played a key role in all of these international developments.

Two of the most hotly debated issues addressed during the revision process concern the issue raised earlier where physical flows and ownership change are not coincident. The two issues are referred to as goods for processing and merchanting. The first deals with cases where goods enter a country for processing but the ownership is not transferred to a domestic producer. In merchanting a merchant buys goods, taking ownership, and sells them to a third party, but the goods never enter the merchant's country.

The Canadian response to these challenges is being developed as part of an overall plan to move the Canadian economic statistics program in line with the revised international manuals mentioned above.

Goods Trade in a Simple World

Goods dominate international trade. To start the exploration of measurement issues, take the simple case of the export of goods across the Canada–US border. There are two firms, one in Canada the other in the US, and one wishes to buy what the other produces. They agree on a contract, specifying the conditions of sale, and in due course the goods arrive at the customs frontier on the way from the seller to the buyer.

In this simple case depicted in Diagram 2 the exporter and the importer of record are the buyer and seller, and the customs documentation would show the flow of goods from one to the other. At the same time, or at least within a short period before or after the shipment, there would be funds transferred from the bank account of the buyer to the bank account of the seller.

Diagram 2: Traditional View of Goods Trade

78
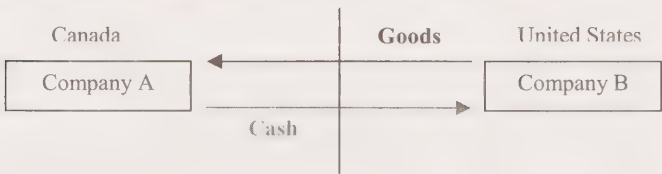
It is a basic tenet of economic statistics such as the national accounts and the balance of payments that transactions should be based on exchanges of economic ownership[3]. In this simple case the goods go from B to A and the money goes from A to B. Since the exact timing of the ownership change between A and B can vary depending on the contract and the payment agreement, as a proxy for change of ownership, economic statistics use the crossing of the customs frontier as the point at which the goods change ownership. It is assumed that this is a good proxy for the change of ownership.

It might be noted that there is no information here on what A is going to do with these goods from B. In the case of Canada and the US we know that often these goods will be used as intermediate inputs in other goods that in turn will go back across the border to the US.

In this simple case, if the industrial activity of each of A and B is known and the location of their business is known, one can build up trade data showing to/from information on a geographic and an industry basis.

The real world, of course, has never been quite that simple, as often transactions are handled via brokers who may affect the timing of certain transactions or at least the recording of transactions. In addition, the timing of payment may vary more widely, and so a debt may be recognized between the supplier and the buyer, which is eliminated as the goods are paid for or delivered if prepaid.

The simple fact of inserting a broker can already cloud the analysis of the data as the customs documentation may well show the industry and location of the broker as one of the transactors and the capacity to undertake industry and geographic analysis is weakened.

The buyers and sellers may be owned by the same owner and thus part of a multinational enterprise. The ownership link is likely to affect the stability of the commercial relationship,

---

[3] The term economic ownership is used to differentiate it from legal ownership, which is generally the same but can differ in cases such as financial leases.

but it may also affect the prices of the transactions that are recorded.

Goods for Processing

The first deviation from this simple model to be explored in this paper is that of goods for processing. As noted earlier, this is not a new phenomenon as examples have existed for decades but there is evidence that the volume of goods traded under this scenario is growing rapidly.

The basic change in the model here is that the goods do not change ownership as they cross the border but the contract between A and B now stipulates that the goods are to remain the property of A but that B will perform some specified processing of the goods and then send them back to A. This type of processing may be between enterprises under common ownership or enterprises operating at arms length. A will pay B a service fee for this processing.

Diagram 3: Goods for Processing – Current Treatment
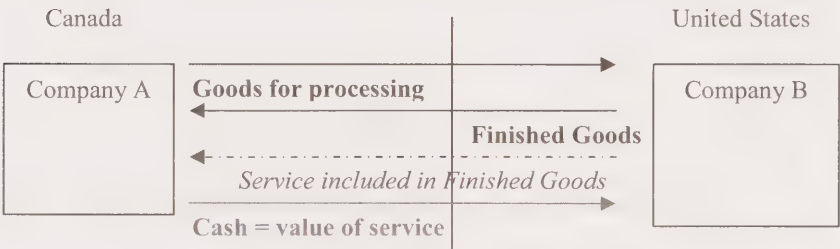
In this scenario the customs process records these goods but they are generally not distinguishable from other goods crossing the border. There must also be a value declared for these goods. However, the counterpart financial transaction between A and B will be greatly different than if ownership had actually changed as the goods moved between them. In this case only the value of the service will flow from A to B, which will correspond approximately to the net difference between the values declared to customs for the import and the export.

The current treatment of these situations in the economic accounts is to record these goods as if they had changed ownership as they cross the border. The full value of the goods is entered into the trade data in both directions and corresponding financial flows are recorded. The service provided by the processor is buried in the value of the returning goods.

This treatment corresponds well to the construction of supply and use tables such as the Canadian I-O Accounts but does not shed light on the evolving behaviour of business activity and is not in keeping with actual financial flows.

During the recent international efforts to update the conceptual guidance for the System of National Accounts and the Balance of Payments, it has been decided that the guidance on these types of transactions will be changed to correspond to the ownership principle. While the goods flowing both ways will continue to be included in the customs data, the convention will be to remove these values from the Balance-of-Payments-based trade data used in the BOP and SNA. Instead the service flow and corresponding payment will be reflected in the accounts.

Diagram 4: Goods for Processing – New International Convention (SNA & BOP)

Canada		United States
Company A	Goods for processing → ← Finished Goods Processing Service →	Company B

Cash = value of service
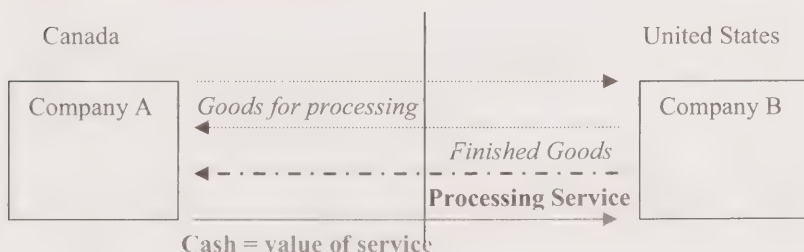
This would require that there be data to differentiate these cross-border flows of goods from traditional transactions. The customs data do not as yet provide such information. The service flows will have to be measured using surveys as there are no administrative sources for these data.

Merchanting .

The new world of production can also have instances of what is referred to in the statistical manuals as "merchanting." Other terms might be international trading or international wholesaling. This is the case where an enterprise in country A buys goods in country B but the goods never enter country A but are sold on to country C.

Diagram 5: Merchanting – International Convention
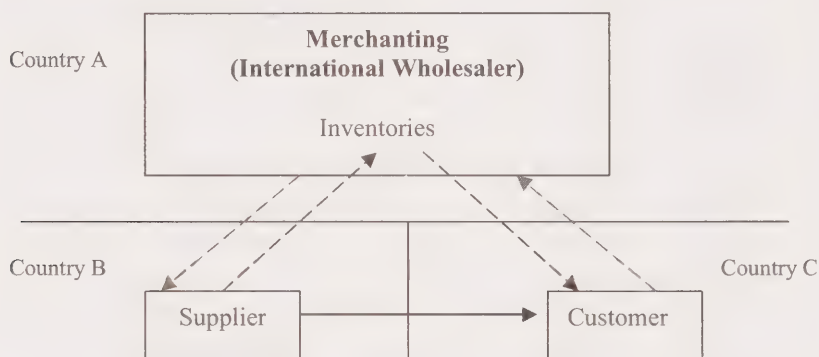
While the ownership of these goods moves from B to A and then from A to C, the customs data will record only a flow from B to C. These supplies will enter into the inventories of the merchant in country A. Surveys of wholesale activity will record these inventory changes and given there have been no imports recorded, the economic accounts will look for domestic production, which of course is not there, to balance out the supply use accounts.

The revised BOP and SNA conventions call for the imputation of flows into and out of country A in line with the ownership changes. It may be practical to develop surveys in country A to gather information from the merchant and may be worthwhile if sufficient activity is present in the country that would distort the signals from these inventories owned outside the country. However, the benefits of collecting additional data on

these transactions for countries B and C are much smaller and it is unlikely that they will adjust their data, thus leading to discrepancies in bilateral trade data.

Goods Production

The goal is to have economic statistics for both production and trade that are consistent and thus provide information to policy makers and Canadians in general on how these aspects of the Canadian economy are developing. Therefore, before turning to look specifically at challenges faced in measuring trade in goods in more fragmented production processes, it may be helpful to review for a moment the traditional way the manufacturing production process is viewed in measuring economic statistics.

Diagram 6 provides a simple case of a producer of goods. The example here has one manufacturing establishment with a separate head office. The head office provides the financing, product development, and marketing, while the manufacturing establishment acquires the other inputs, manages the inventories, and produces finished goods for market.

In the model, please note that the activities/outputs of the head office are considered ancillary services proved to the manufacturing establishment.

Diagram 6: Traditional View of Goods Production

83
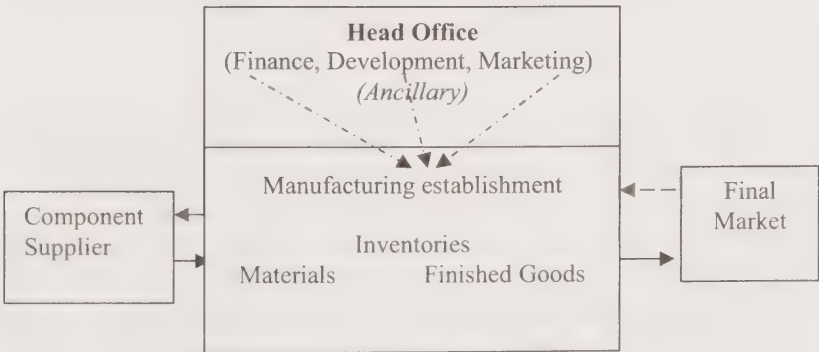
The model can explicitly be extended to include trade in goods and services as in Diagram 7. Note that if some of the services provided by the head office in the previous example are outsourced, then they are still treated as being supplied through to the manufacturing establishment.

Diagram 7: Traditional View of Goods Production with Trade in Goods and Services
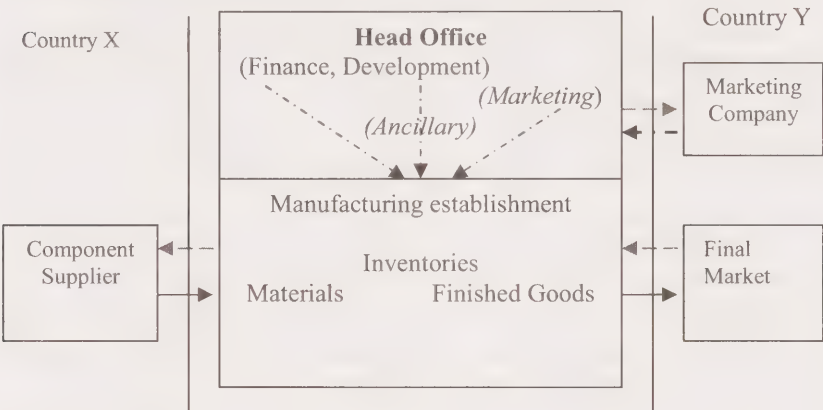
Outsourcing of the Goods Production

Diagram 8 provides an example with outsourcing of the procurement of input materials and the manufacturing process. In this case, all of the goods produced are "exported" to a country other than that of the head office or manufacturing plant. One could think of this as a case of line of products produced only for an export market.

In this example the head office and the manufacturing plant are no longer in the same economy and are two separate enterprises. The head office still produces the financing, product development, and marketing but acquires all other inputs from outside the country. In this case, the head office enterprise buys all of the material inputs and has them sent to the establishment of the processor to be assembled. The head office maintains ownership of these materials as in the goods for processing case discussed earlier. Therefore, while processing may all take place outside of the country of the head office, it is the head office that

bears the financial risks associated with carrying these inventories throughout the production process. This case also has elements of the merchanting case discussed earlier, as the head office acquired the ownership of these materials but they never physically enter the country, thus never appear in the customs data.

There are two important differences from the traditional view of the producing enterprise that should be noted:

- Inventories of materials and finished goods have moved from the manufacturing establishment to the head office.
- The services produced by the head office – finance, product development, and marketing – are no longer supplied to the manufacturing establishment.

In Diagram 8 the flows of goods and payments are indicated. It is clear immediately that the path followed by the cash to pay for these transactions does not correspond to the path followed by the goods themselves. Again this is similar to the merchanting example.

Diagram 8: Outsourcing Production for Export market – Flows of Goods and Cash
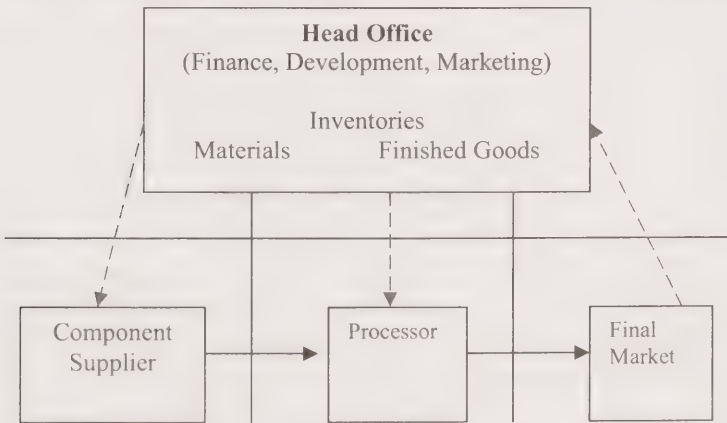
Using the principles noted earlier about ownership as a basis for recording, Diagram 8 indicates how economic data might be recorded to reflect the economic activities rather than the physical supply-use activities of this production process.

The purchase of materials would be shown as a transaction between the head office and the component supplier. The processing services would be supplied to the head office. The final sales of finished goods would be transactions between the head office and the final market economy.

The implication of this presentation of the economic data is that the head office, even though it has no manufacturing facilities or warehousing facilities, is the producer of the finished goods. While this does not correspond to the physical flow of the goods, it does recognize the behaviour of the economic agents in this situation. It is the head office that is undertaking the economic decisions that govern production, whereas the processor is offering a specific service only.

Ideally, to be consistent with this recording of production and trade, the bottom three countries in Diagram 9 would remove the goods flows from supplier to processor and on to final markets from each of their accounts[4]. However, as with the merchanting case described earlier, the information to undertake such adjustments may be difficult to obtain.

Diagram 9: Outsourcing Production for Export market – Economic Flows

---
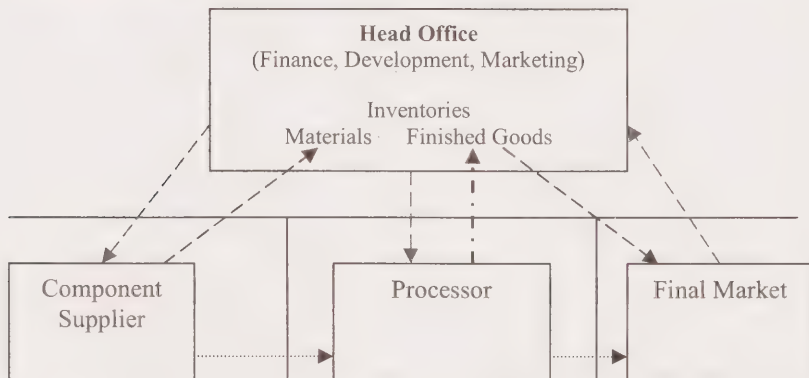
[4] Information on physical flows will still be important for economic development as infrastructure requirements are determined by these physical flows.

Diagram 9 shows four different countries. Canada can have enterprises that correspond to those shown in each of the countries. That is, we face the measurement challenges depicted in each of the four fictional countries. This section will take a brief look at the challenges in measuring the activity of each of the four enterprises.

Starting with the component supplier, the challenge is the geographical allocation of trade. The customs data will indicate the country of the processor while if we ask the supplier, the response would be the head office country. Surveys of goods producers such as this supplier ask for little on the geographic distribution of sales of goods, particularly for sales outside the country. This is, of course, because the customs data are used to provide this information. While value-added data will be correct, the information on trade patterns available to trade policy analysts and negotiators will be affected by how these transactions are measured.

Turning to the country of the processor, the information on the transactions between the processor and the head office will have to be collected by survey. Up to now in Canada these cases have been handled on a case-by-case basis, with special reporting arrangements set up with processors. It is not possible to identify the customs records associated with this processing activity and it is unlikely that this would change in the near future. If this activity is limited to a few large players, this case-by-case treatment can be used, but if it becomes (or already is?) more widespread, then this approach is too costly to be applied to a large number of cases.

The case of the final market economy is similar to that of the supplier country, in that the geographical allocation of supply based on the customs data and survey data would differ. Surveys of wholesale and retail firms importing goods do not ask for details on the geography of supply. There may also be a valuation difference if the value of imports declared for customs purposes does not correspond to the full value of the purchase price paid to the head office. Reconciliation of these values will

only occur at a macro level in constructing the economic accounts.

The large challenge with measuring the economic activity of the head office is that surveys must be used to collect a lot of detail on the goods and service inputs and outputs of this enterprise. Collecting detailed commodity and geographical information using surveys is very burdensome on the respondent. However without sufficient detail on these commodities, it will be difficult to distinguish the traditional manufacturing operations from those which have outsourced significant parts of the production process.

Country of Origin and Ownership

Customs officials collect data on the "country of origin" of the product, which is based on customs rules for imports and the country of destination for exports. Also collected is the point of shipment. This transfer process also means that the shipper (exporter) may not know the ultimate destination of the goods, whereas the importer generally knows the country of origin due to the rules of origin.

This can result in significant differences in the bilateral trade figures for some country pairs. For example, Mexico reports about twice the value of imports from Canada as we report exports. The difference is largely goods that the exporter declares as going to the US but that are just passing through.

The case is similar for Chinese goods entering Canada. About half of the Chinese imports on a country-of-origin basis arrive in Canada having a point of shipping of either Hong Kong or the US.

If ownership is to be a fundamental basis for economic statistics, the question arises as to which of the geographical data on the customs record – country of origin or country of shipment – is the best indicator of the counterparty to the transaction. Are the Chinese goods transiting through the US on their way to Canada simply just 'in transit' or are they being acquired by an entity in the US which then distributes them (sells them) to a network of North American outlets?

Services trade is still the poor cousin but, as seen in earlier examples, fragmentation and reorganization of the production chain may be leading to significant increases in the relative shares of services trade. In the 'good old days' as it were, services were less of a concern. It was assumed that most services had to be produced and consumed at the same time so that the international trade in services was limited.

The ease with which producers can travel to other parts of the world and the advances in electronic delivery of many products has certainly changed the potential for international trade in services. In simple volume terms it is still much smaller than goods trade but services trade has seen much greater rates of growth for a number of years now.

The other point that is likely clear by now is that there is a growing fussiness about the differentiation of what a good is and what a service is. Thus, any exercise that is forward looking should cover both.

There are of course no customs documents for services[5]. The data for services trade generally come from business surveys. This measurement approach has very different strengths and weaknesses from the administrative process (customs) used for goods.

In general, businesses can more readily respond to surveys on their sales rather than purchases of inputs, particularly if one is asking about the geography of the transaction and the industrial activity of the counterparty. Whereas for goods the import data are generally of better quality, as customs administrations are more diligent about collecting duties owed and due more recently to security concerns, trade in services is better at measuring exports.

The use of surveys also limits the amount of detail on geography and commodity detail for services transacted, since asking for very specific detail on the service provided and the loca-

---

[5] Customs documents are used to develop estimates of transportation services for the delivery of internationally traded goods.

tion of the counterparty rapidly becomes a very large burden for the respondent.

Cross-Border Valuation and Transfer Pricing

While a significant proportion of international trade has for some time been between affiliated enterprises, it is difficult to identify these trade flows separately due to the use of brokers and other agents. It has been recognized that intra-firm transactions may reflect transfer pricing and thus affect the valuation of economic activity between economies.

The customs 'transactions' depicted in Diagram 3, even if they are with non-affiliated parties, may not reflect arms-length measures of economic valuation as the actual transaction taking place is for a service but the valuations that are being declared, and currently used in economic accounts, are for the goods crossing the border. It is generally assumed that the difference between the value of the goods entering for processing and those returning are equal to the contracted service from the processor.

The customs flows in the scenario in Diagram 9 are also displaced from the actual economic transactions with the customs values both entering and exiting the country of the processor, presumably based on the accounts of the head office.

If values declared for customs purposes are not the same as the actual transaction values then the difference will distort measures of economic activity.

Structural vs. Behaviour-Based Economic Data

A complete picture of the Canadian economy requires data on both the structural aspects of our economy and the behavioural aspects. The national accounts address the need for structural data primarily via the supply-use tables and the structural data published by individual surveys such as the annual survey of manufacturing. The behavioural data is primarily delivered via the sector accounts of the national accounts and related sector-specific data from programs such as those in the Balance of Payments and Public Institutions Divisions.

There is evidence that the organization of production and trade is changing. The challenge for the statistical system in part is to decide when these changes are sufficiently important to change the basic focus of the accounts and underlying surveys. To refocus the statistical system and move away from the traditional models used as the foundation of the economic data would be very costly.

The costs to adjust to a different focus are not only those that the statistical agency will have to bear but also those imposed on respondents if surveys are expanded. Significant changes to the model underlying data production may also cause discontinuities in some data series that will challenge analysts' capacity to understand the evolution of the economy. Of course remaining with the current model may obscure structural changes underway in the economy raising challenges for analysts.

Conclusion

There are clearly cases of each of the different models of economic activity discussed in this paper in the Canadian economy. What is less clear is the proportion of activity that falls into these different scenarios. The challenge for the statistical system is to find ways to measure these emerging phenomena in a manner that is cost efficient in terms of both response burden and budget. Given resource constraints, which activities—for example, goods for processing, merchanting, trade services— should perceive priority for development?

Consultation with policy makers, business respondents, academics, and other analysts will be important in mapping a way forward to address these important challenges. At the same time the statistical system needs to keep in step with international statistical developments if Canada is to have measures of economic activity that are comparable across countries.

Annex
Statistics Canada Programs Addressing Globalization

Statistics Canada has several longstanding programs that provide information on globalization issues . The balance of payments is the macroeconomic core of these measures with specific data available for key supporting programs such as international trade in goods, international trade in services and foreign direct investment. In addition, Statistics Canada has administered the Corporations Returns Act since its inception.

Corporations Returns Act

The original Corporations and Labour Unions Returns Act (CALURA) was introduced in 1962. Parliament amended the Act effective January 1999, changing it to the Corporations Returns Act (CRA) and removing the requirement for labour unions to report.

Under the CRA, Statistics Canada prepares an annual report to Parliament on the foreign ownership and control of Canadian businesses that examines financial and ownership information on corporations conducting business in Canada. This information is used to evaluate the extent and effect of non-resident control of the Canadian corporate economy. The CRA data are a primary source of information about mergers and acquisitions, foreign control of enterprises, corporate concentration and the legal structure of enterprises in Canada.

The ownership information collected from the returns filed by Canadian corporations under the Corporations Returns Act is publicly available by law and Statistics Canada uses the information to compile an inter-corporate ownership directory showing "which corporation owns which other corporation" in Canada. The directory tracks the ownership of the largest Canadian corporations and provides up-to-date information reflecting recent corporate takeovers and other substantial changes. Ultimate corporate control is determined through a careful study of holdings by corporations, the effects of options, insider holdings, convertible shares and interlocking directorships.

Finally, the data gathered under the Corporations Returns Act have been used to do further analysis on topics such as: the impact of mergers and acquisitions on corporate profits; changes in foreign control under different regulatory climates; the evolution of foreign bank subsidiaries and full-service branches in Canada; mergers and acquisitions and their relationship to foreign control; and foreign control and corporate concentration.

Foreign Direct Investment

Foreign direct investment has been an important source of capital for the Canadian economy for many years. In recent years Canada has also become an important supplier of foreign direct investment to other countries. Currently the net stock of assets resulting from foreign direct investment activities provides a positive contribution to Canada's net wealth, although the balance on portfolio investment continues to be negative.

The latest release of FDI position statistics showed outward FDI positions in around 150 countries. For direct investment coming from abroad (so-called inward FDI) the detailed statistics show investment positions from about 100 countries. Industry detail is also available. Data on FDI transactions and income are available with considerably less country detail (6 regions or countries). Confidentiality requirements constrain Statistics Canada's ability to release detailed FDI flow information, although the increased FDI activity in recent years and organizational and methodological changes within Statistics Canada may mean that additional flows detail and new indicators such as FDI statistics by country of control may be released in future.

Trade in Services

The increased specialization of production is leading to a greater importance for trade in services, which is relatively difficult to measure. Statistics Canada recently embarked on a three-year project to improve the data for international trade in services.

There are two main objectives for this project. The first aims at improving the accuracy of the annual estimates of total imports and total exports of commercial services within total trade in services. The population of Canadian businesses that are involved in international services transactions needs to be better identified, particularly with respect to small and medium sized businesses. Therefore, improved identification of the target universe is an important part of the business objective.

The second objective is to improve the commodity, geography and industry detail for trade in services estimates.

Commodity detail will be improved by collecting services commodity categories that map easily into the North American Product Classification System (NAPCS), the Central Product Classification (CPC) and the Extended Balance of Payments Service Classification (EBOPS). If it proves feasible, the number of services commodity categories collected on the annual trade in services survey will be increased. This survey currently collects international trade data for 32 services commodity categories.

With respect to geographical detail, a provincial breakdown of Canada's imports and exports of services by commodity will be developed on an annual basis. The project will also explore ways in which Statistics Canada's Balance of Payments program can respond to changing economic conditions and user requirements by modifying, as needed, the partner country aggregations and detail that are published. This approach aims at providing as much useful information as possible, while simultaneously respecting the need to maintain data confidentiality for individual enterprises and limiting the survey response burden.

In connection with the efforts to measure international trade in services more effectively, as part of its redesign of its annual program, Statistics Canada's services industries program is developing a new module that will allow for the collection of additional data on imports and exports of services in selected industries, at the same time as other services financial statistics are collected.

Trade in Intellectual Property, Licenses for Patents, Trademarks and Copyrights

Statistics Canada collects data on the technological balance of payments that include payments and receipts for the acquisition and use of patents, licenses, trademarks, designs, know-how and closely associated technical services, as well as for research and development services. The data derive from the Survey of Research and Development in Canadian Industry. The data on R&D services are reconciled with balance of payments data in order to find firms that purchase R&D services but perform no R&D themselves.

Trade in Goods: Exporter/Importer Registers

The international trade statistics program has produced an Exporter Register database which provides reliable counts of exporting establishments and their value of merchandise exports over the 1993 to 2004 period. An Importer Register database, currently under development, will provide similar data for Canadian importer establishments. Currently, preliminary importer data are available for 2002.

The Exporter Register provides an invaluable longitudinal database on the characteristics, performance and evolution of Canadian exporters. This information is delineated over several dimensions including industrial classification (NAICS), exporter size, destination of exports and province of residence of the exporter. This database provides counts of establishments exporting merchandise and the value of their exports by employment size category beginning with reference year 2000.

The Exporter Registry provides statistics on the characteristics of exporting firms. This is allowing research to be conducted on important policy issues such as the determinants of export success, and the relationship between exporting and productivity. This database provides empirical evidence to aid in the evaluation of many programs aimed at supporting and assisting exporters.

The Importer Register database will provide similar data for Canadian importer establishments. As noted, data are cur-

rently available for 2002 and data for 2003-2005 will be available in early 2007.

The union of the Exporter and Importer databases will contribute significantly to current research being conducted on issues such effects of recent exchange rate changes and other topics related to globalization.

Foreign Affiliate Trade Statistics

The delivery of services to international markets is often accomplished through foreign operations or foreign affiliates. Starting with data for 1999, Statistics Canada has an annual program that provides data on the sales and employment of these foreign affiliates of Canadian firms for both the goods and services sectors. This program shows that this delivery mode is more important than cross-border delivery of services.

Recent Developments and Future Challenges

More recently, Statistics Canada has taken decisions to reallocate resources to programs that can address data requirements in the general domain of globalization. The most important are the efforts to improve statistics for services output and productivity. These improvements can be grouped into three categories: (i) improved coverage of core industry data through an expanded annual services industries survey program; (ii) improvements to sub-annual indicators of economic activity and (iii) an expanded program of services price indexes. This note does not cover these, but descriptions are available in other documents.

This section describes some additional initiatives that have started very recently or are in the more advanced stages of development.

Globalization Project

In response to the data demands of policy makers and the new statistics proposed in the OECD Handbook mentioned earlier, Statistics Canada has allocated funds to a globalization indicators project. This project is focussing mainly on developing basic infrastructure for the production of additional globalization

related data series. Subsequently, this new infrastructure will permit the development of a suite of globalization indicators.

An important feature of this project will be the identification of Canadian multinational enterprises (MNEs). While it has been possible to identify the operations of foreign multinational operations in Canada for some time using the Corporations Returns Act, it is only with the recent linking of the foreign direct investment program to the business register that Canadian multinationals can be easily identified.

Business Register

Statistics Canada's business register is presently under redevelopment and as part of this effort additional emphasis is being put on recording the international links between Canadian businesses and their foreign parents or subsidiaries. Key characteristics of these linkages will be recorded to support future data development and research.

Several data sources will be tabulated by ownership class (foreign MNE, Canadian MNE, and non-MNE) and by trade status (importers, exporters, both, or neither) to indicate the level of globalization of the Canadian economy.

1. Principal statistics from the annual Survey of Manufactures will be tabulated by ownership class and trade status for reference years 2000 through 2004 (2005 where available). Statistical tables will be produced by industry and province. The principal industrial statistics include shipments, employment, salaries and wages, cost of materials and supplies used, cost of energy used, goods purchased for resale and manufacturing value added.

2. Investment expenditures data (both construction and machinery and equipment) will be produced by ownership class and by trade status for reference years 2000 through 2005 (2006 where available). Estimates will be prepared by industry and province.

3. Statistics on research and development expenditures in Canadian industry will be produced by ownership class and trade status for reference years 2000 through 2004 (2005

where available). Variables such as sources of funds for intramural R&D and value and type of intramural expenditures will be included.

4. In addition, results from the market and supply chain questions included on the 2005 Innovation Survey, which targets the manufacturing and logging industries, will be tabulated and analyzed. The analysis will explore variations by ownership class and trade status, and by industry and province.

Global Value Chains Data Development – Feasibility Study

A feasibility study was recently approved by the Policy Research Data Group that will present and examine different options for collecting data on the phenomena of global value chains, and recommend a course of action for so doing. The study proposes to do the following things:

1. Clarify concepts and definitions. The U.S. Congressional Government Accountability Office developed a conceptual framework that defined offshoring in relation to other related concepts such as foreign direct investment and trade. The OECD has adopted this conceptual framework and several countries have used it to develop surveys and perform research. There is, however, some work to be done for the implementation of these concepts in survey-taking. This means certain concepts associated with offshoring need to be related to concrete and measurable activities and characteristics. First, the business functions that are offshored need to be classified in a meaningful way (e.g. legal functions, payroll functions). Second, the characteristics of those functions need to be classified (e.g. knowledge intensive, high wage, capital intensive, intellectual property related). Third, the forms that the offshoring relationships take must be classified (e.g. contracts, joint ventures, equity).

2. Review the results from other countries' surveys on offshoring, as well as Canadian surveys related to the matter. This is important background for the development of options for the subsequent steps. With respect to the Canadian situation,

98

the intent is to build upon the expertise gained in the following surveys which had questions related to offshoring: 2005 Innovation Survey, Natural Health Products Survey, Language Industry Survey, and the functional Foods and Nutraceuticals Survey.

3. Describe and recommend the survey strategy. The strategic elements include the choice of observation unit and the proposed survey frame, and the description and field-testing of a survey vehicle. The final sample size will be dependent on the amount of detail expected to be derived from the sample (geography, firm size, firm ownership and industrial detail).

Commercialization

Funding was also approved by the Policy Research Data Group to undertake a pilot survey of commercialization activities in Canada. This survey will examine how new developments are brought to market, both for projects undertaken in Canada and commercialized here and for projects developed by Canadian firms outside the country to exploit new market opportunities. The survey will be conducted in 2007 for reference year 2006.

International Labour Supply and Remittances

Globalization is not just a business enterprise phenomenon. With an increasing proportion of the labour supply in many countries coming from temporary and permanent migration, households and extended families are becoming more global.

The large immigrant populations in many countries, including Canada, transfer significant amounts of resources to their home countries to support family members that remain behind. The importance of developing reliable estimates of these international remittances by persons has been recognized by international financial and statistical organizations, as well as by the leaders of the G8 countries at the Sea Island Summit in July 2004. Canadian measures of these flows are poor and the possibility of improving these data is being reviewed.

The Survey of Household Spending (SHS) collects some data on this phenomenon, but the variance is high given that the sample is not designed to ensure adequate representation of the

immigrant population for this particular purpose. Furthermore, the SHS does not provide any breakdown of remittance payments by country of destination. It may prove feasible to improve Canadian remittance statistics either through changes to existing survey vehicles or as part of a new household survey. Some relevant changes are planned for the Survey of Labour and Income Dynamics (SLID), including questions about household-to-household transfer payments.

The Longitudinal Survey of Immigrants does pose questions related to international household-to-household transfers. However, a major restriction is that the target population includes only the immigrant people who arrived in Canada between October 2000 and September 2001, which is not representative of Canada's immigrant population as a whole for this particular purpose. Nonetheless, it may be possible to derive some useful conclusions.

Whereas most industrialized countries produce separate estimates for employment income transactions with non-residents, Canada does not. A project to improve statistics regarding international compensation of employees was launched in early 2006. The project is exploiting administrative data. Updated statistics on earnings of foreign residents employed by Canadian employers are expected later in 2006. It may not prove possible to improve statistics on the earnings of Canadians employed by foreign employers via domestic administrative data sources. Therefore, an exchange of statistics with partner countries will be examined. An additional benefit is that some results of this project will likely be useful for improving estimates of other components of the balance of payments such as remittance statistics and a special class of trade in services statistics, identified as "mode 4" in the General Agreement on Trade in Services (GATS), where a service supplier of one country travels to another country to deliver the service.

References

International Monetary Fund. 1993. *Balance of Payments Manual.* Fifth Edition, 1993. Washington, DC. International Monetary Fund.

United Nations, Commission of the European Communities, International Monetary Fund, Organization for Economic Co-operation and Development, and World Bank. 1993. *System of national accounts 1993*. Series F, no. 2, rev 4. New York: United Nations.

Wilson, Karen. 2006. "The Architecture of the Systems of National Accounts, A Three-Way International Comparison of Canada, Australia and the United Kingdom". In *A New Architecture for the U.S. National Accounts*, ed. Dale W. Jorgenson, J. Steven Landefeld, and William D. Nordhaus, 113-142. Chicago: The University of Chicago Press.

Trade Theory, Trade Policy, and Cross-Border Integration

Michael Hart and Bill Dymond[*]

Introduction

In a 1992 article, economist Paul Krugman asked whether the new trade theory required a new trade policy. He concluded that it did not, arguing that since existing trade policies were unrelated to long-established trade theories[1], it was difficult to see why such policies needed to accommodate new theoretical insights. In his view, "GATT-think [is] a simple set of principles that is entirely consistent, explains most of what goes on in the negotiations, but makes no sense in economic terms."[2] Krugman's point was clever but also misleading because it did not fully appreciate either the nature or the genius of the international trade regime, or the challenges it needed to address in the

[*] Hart is Simon Reisman Chair in Trade Policy at the Norman Paterson School of International Affairs and a distinguished fellow of the Centre for Trade Policy and Law at Carleton University. Dymond is senior executive fellow at the Centre. This paper was originally prepared for the conference *Integrative Trade between Canada and the United States Policy Implications*, organized by the Centre For Trade Policy and Law, held at Ottawa University, December 6, 2006. The views expressed are those of the authors and are not to be attributed to Foreign Affairs and International Trade Canada or to the Government of Canada.

[1] Theories first developed by, *inter alia*, Adam Smith and David Ricardo (specialization and comparative advantage) and refined over the years by economists like Eli Heckscher, Bertil Ohlin, and Paul Samuelson (the factor proportions theory) and Ray Vernon (the product cycle theory), and now supplemented by theories to take account of imperfect competition, increasing returns to scale, and other factors.

[2] "Does the New Trade Theory Require a New Trade Policy?" *The World Economy* 15:4 (July 1992), p. 429. See Douglas Irwin, *Against the Tide: An Intellectual History of Free Trade* (Princeton: Princeton University Press, 1996) for a full discussion of the development of trade theories and the broad commitment of economists to the doctrine of free trade.

face of deepening global integration. The system of rules and procedures worked out and applied in the period from the founding of the US Reciprocal Trade Agreements program in 1934 through the Uruguay Round of GATT negotiations in the 1990s embodied more sensible economic content than Krugman cared to admit. At the same time, he seemed to discount the extent to which there needed to be new thinking about trade policy, some of it in response to new trade theories developed over the past few decades, but even more to take account of the changing nature of international production and exchange.

The new thinking needs to learn from insights developed in industrial organization theory, economic geography, business economics, and other domestically oriented sub-branches of micro-economics, suggesting that there are fewer differences between international trade and domestic commerce than was long assumed. As Krugman himself explained in a later article, "the trend in manufacturing has been to slice up the value chain—to produce a good in a number of locations, adding a little bit of value at each stage."[3] This kind of fragmentation was originally confined to the firm and then to spatially proximate and, often, related firms. Increasingly, however, neither geography nor ownership remain as serious obstacles to the fragmentation of production and its subsequent integration. Even more pertinent to international trade, national borders are no longer significant barriers to the organization of production. Insights into the nature of production and the operations of private firms have opened up new ways of explaining international trade and investment. They have also suggested the need for new thinking about the object and purpose of trade and related negotiations among governments.

The old trade policy, despite Krugman's dismissal of GATT-think, was firmly grounded on the theory of comparative advantage. It assumed that trade is largely a matter of exchanging goods, that it takes place between national economies, reflecting each country's resource endowments and comparative

[3] Paul Krugman, "Growing World Trade: Causes and Consequences." *Brookings Papers on Economic Activity*, 1 (1995), p. 334.

advantage. It further implicitly assumed that national production was associated with the use of mostly domestic resources, meaning that export expansion was associated with a commensurate increase in domestic value-added, jobs and profits. The old trade policy sought to remove national barriers to trade in goods on a progressive, i.e., politically sustainable, basis and to make the rules of the game fair, transparent, and non-discriminatory. Those were the means; the objective was to let markets work in order to take advantage of specialization and thus contribute to growth in national and global welfare. It worked: among OECD economies, trade grew at a rate that far outstripped growth in production to the point that an increasing share of national welfare was being derived from international exchange. Interestingly, as international trade and investment flourished, more and more could not be explained solely on the basis of comparative advantage and national endowments. Economists responded with a range of more sophisticated explanations, including insights derived from analysis of the domestic economy[4].

Of course, Krugman was reluctant to admit that the global economy was beginning to operate more and more like the large, continental US economy of the early post-war years. In the 1990s, he popularized sophisticated economic arguments to demonstrate the extent to which production in the United States was still largely focused on goods and services consumed at home[5]. In terms of a traditional reading of trade statistics, he was right, but in pointing out that some of the economic problems being experienced in the United States at that time were

---

[4] See, for example, Patricia Dillon, James Lehman, and Thomas D. Willett, "Assessing the Usefulness of International Trade Theory for Policy Analysis," in John S. Odell and Thomas D. Willett, eds., *International Trade Policies: Gains from Exchange between Economics and Political Science* (Ann Arbor: University of Michigan Press, 1990), pp. 21-54 for discussion of some of the newer trade theories.

[5] See, for example, "Competitiveness: A Dangerous Delusion," *Foreign Affairs*, vol. 73, no. 2 (1994); 'Does Third-World Growth Hurt First World Prosperity?' *Harvard Business Review*, July-August 1994; and with Robert Lawrence, 'Trade, Jobs and Wages,' *Scientific American*, April 1994.

homegrown, he missed another point: traditional ways of inter-preting trade statistics no longer captured the extent of international economic interdependence. As Douglas Irwin points out, "A close analysis of the merchandise trade figures indicates that trade is substantially more important now than in the recent past for those sectors engaged in trade."[6] He calculated that US merchandise exports as a share of merchandise production grew from 15 percent in 1970 to nearly 40 percent in 1999, even though the share of merchandise trade to GDP grew much more modestly, largely because of the growth of service production as a share of GDP. In Canada, the exports of goods in 1999 represented 125 percent of the value of goods production, consistent with the high level of imports in Canadian exports and the much more export-intensive nature of production in Canada. The comparable figure for 1970 was 65 percent, suggesting a similar rise in the export intensity of the economy[7].

Cross-border fragmentation and integration of production between Canada and the United States started well before it had become commonplace on a global scale. As a result of the 1965 Autopact, automotive production in the two countries was re-organized in the 1960s to allow the major assemblers and their suppliers to integrate their facilities on both sides of the border and thus provide scope for more specialization and the benefits that flow from economies of scale and larger markets. As border barriers came down, other industries followed suit, a process that deepened and accelerated following the implementation of the 1989 Canada-US Free Trade Agreement (CUSFTA) and the 1994 North American Free Trade Agreement (NAFTA).

In this paper, we consider the implications of deepening and accelerating cross-border integration for North American trade policy. We argue that the benefits that flowed to the two economies from the conventional trade negotiations of the past have now been largely realized. More benefits have flowed

---

[6] Douglas A. Irwin, *Free Trade Under Fire*, 2nd edition (Princeton: Princeton University Press, 2005), p. 8.

[7] Values calculated from Statistics Canada GDP and Balance of Payments data.

from the more innovative provisions in the CUSFTA and NAFTA. Nevertheless, there remain barriers to the full exploitation of cross-border integration. To achieve these additional benefits, a concerted effort will be required to address the combined impact of dated, dysfunctional, and intrusive border administration, the haphazard process leading to deepening regulatory convergence, and the frail institutional capacity to govern accelerating integration, together with an expansion of the reach of binational policy-making to address investment, intellectual property rights, labour, services, and other economic transactions.

Thus, while the trade policy of the past may have reached the point of diminishing returns, new challenges have arisen that we believe to be amenable to resolution through bilateral negotiations. Progress on these issues requires a better understanding of the successes and limitations of the old trade policy, the contribution of regional trade negotiations, the nature of modern production and exchange, their cross-border manifestation between Canada and the United, and the barriers to their full development. Following a discussion of these factors, we will consider the new trade policy issues that need to be addressed between Canada and the United States and the benefits that should flow from their successful resolution.

The Old Trade Policy

Over a period of some fifty years—from its first deployment in the US Reciprocal Trade Agreements Program in the 1930s through the conclusion of the Tokyo Round of multilateral trade negotiations in the 1970s and into the opening of the Uruguay Round in the 1980s—trade negotiations grounded in mercantilist bargaining and negative prescription proved an unqualified success for those countries prepared to take advantage of this strategy. The result provided the basis for a tremendous increase in international trade and a major boost in productivity and prosperity. Mercantilist bargaining offered a politically acceptable way to pursue politically difficult economic goals.

107

Through such trade negotiations, governments sought to increase economic welfare by reducing discrimination, removing barriers, disciplining potentially distorting policy measures, and providing greater scope for the operation of the market, unencumbered by artificial policy barriers. They also sought the order and stability that come with clear rules and equitable dispute settlement provisions. In short, both legal and economic maxims suggested that international trade agreements were mutually beneficial. Every barrier removed or rule established would benefit all participants.

Such trade negotiations, however, were in fact pursued on the basis of pre-economic and pre-legal concepts: every "concession" granted by one party had to be matched by a "concession" from another. Ministers and their officials, while they knew better, behaved as if they were engaged in a zero-sum game. Reason told them that the removal of barriers and the establishment of rules made sense, but politics dictated that they could only achieve these goals by seeking greater export opportunities while minimizing the prospect and benefits of increased import competition. Political discourse was based on the conceit that the strength of the nation required a positive trade balance, an ability to do without imports, and the promotion of strategic advantage over all other nations. British economist Martin Wolf first called this process mercantilist bargaining and suggested that, perversely, it was only through mercantilist bargaining that progress could be made[8].

Mercantilist bargaining also proved superbly suited to the architecture of the US-inspired reciprocal trade agreement: an interlocking set of negative prescriptions by which governments undertook self-denying ordinances disciplining their capacity to impose trade barriers at the border and to discriminate among domestic and imported goods or among their trading partners.

---

[8] See Martin Wolf, "A European Perspective," in Robert M. Stern, Philip H. Trezise and John Whalley, eds., *Perspectives on a U.S.-Canadian Free Trade Agreement* (Washington: Brookings Institution, 1987). In *Why Globalization Works* (New Haven: Yale University Press, 2004), he provides a much more complete picture of the role of trade negotiations in spreading the benefits of globalization.

Its substantive obligations required participants to refrain from applying their trade policies in ways that were contrary to these fundamental rules. The obligations did not require adherence to absolute standards of behaviour, nor did they impose rules and procedures for the detailed administration of trade policy. There was no obligation in the GATT, for example, to set maximum tariff rates. GATT members remained free to negotiate such rates and free to apply them at lower levels or not all. The basic national treatment obligation did not guarantee a standard of treatment for imports, for example, respecting the level of indirect taxes, but rather required that such taxes or other regulations affecting internal trade not be higher or more burdensome than those applied to products of domestic origin.

With respect to non-tariff barriers, the GATT did not require its members to apply countervailing or antidumping measures, marks of origin, quantitative restrictions on trade, or subsidies, but rather established disciplines on the use of such measures. The articles allowing exceptions for balance-of-payment reasons, tariff preferences, or import surges were similarly encumbered with disciplines and in some cases surveillance, all intended to make their use difficult. While the articles on customs administration and valuation set out prescriptive rules and procedures, the need for such measures arose from the application of an import regime which the GATT did not substantively require[9]. The arch over all these obligations was the general requirement of most-favoured-nation treatment.

The focus of these negotiations was barriers to the exchange of goods. Barriers to the flows of services, investment capital, the internationalization of intellectual property rights, temporary business travel, labour, and other cross-border transactions were governed, if at all, by separate international instruments, none of which included the kinds of rights and obli-

---

[9] Article X on the publication and administration of trade regulations is an exception to the paradigm of negative prescription. It should be added, however, that the existence of a body of norms (rather than requirements), did lead to a gradual convergence in the trade policies of the core GATT members, the industrialized countries clustered around the North Atlantic.

gations that became the basis of the trade-in-goods regime. International air services, for example, were governed on the basis of a series of interlocking bilateral arrangements that were crafted not only on the basis of mercantilist bargaining, but pursued in order to achieve mercantilist results. Unlike the equality of opportunity that was critical to the goods regime, the patchwork of other international arrangements was much more geared to specific outcomes. The intellectual property regime embedded in the arrangements administered in the World Intellectual Property Organization (WIPO) relied on comity rather than enforceable contract.

In the six rounds of multilateral trade negotiations between the provisional entry into force of the GATT in 1948 and the completion of the Tokyo Round in 1979, there was no significant change in this fundamental architecture. The focus of each negotiation was the reduction of tariffs on industrial goods through mercantilist bargaining. While the Kennedy and Tokyo Rounds expanded negotiations beyond tariffs to the development of disciplines on export and internal subsidies, the application of countervailing and antidumping duties, and separate agreements on trade in civil aircraft and government procurement, the results were fully consistent with the principle of negative prescription[10]. In two areas of the Tokyo Round, however, there was already evidence of the pending shift in the centre of gravity from negative prescription to positive rule-making. Agreements governing customs valuation and technical barriers made a tentative start at requiring adherents to administer regulations along specified lines. Neither agreement, however, required countries to impose a tariff on imports or to maintain a products standards regime.

The Uruguay Round fundamentally altered the nature of the GATT-based trade relations system[11]. In addition to some ma-

---

[10] For an analysis of the Tokyo Round, its origins, its objectives, issues and results see Gilbert R. Winham, *International Trade and the Tokyo Round Negotiation* (Princeton: Princeton University Press, 1986).

[11] For a useful summary, see Jeffrey J. Schott, *The Uruguay Round: An Assessment* (Washington: Institute for International Economics, 1994). See

jor achievements within the traditional framework, it made a decisive shift toward positive rule-making as the basis for a new architecture of international trade rules. It also sought to extend the ambit of the rules to other economic transactions, particularly trade in services, some trade-related aspects of investment, and intellectual property rights. The shift into rule-making is particularly apparent in the agreements governing trade in services, the protection of intellectual property, technical barriers, and sanitary and phytosanitary measures, and in parts of the agriculture agreement[12]. When combined with the positive obligation to ensure conformity with WTO rules, a more potent dispute settlement system, the trade policy review mechanism, and the new disciplines on services and intellectual property rights, the WTO institutionalized a degree and intensity of intervention into domestic governance which exceeds anything possible or contemplated under the GATT.

The stalemate at the Doha Round of negotiations suggests that the transition from a regime of negative to positive rules was not fully thought through. The bargaining techniques that had worked so well for more than fifty years proved less well suited to the emerging architecture and the much more comprehensive ambit of the rules. As discussed further below, regional negotiations relied less on mercantilist bargaining and proved more adept at dealing with the politically tricky issues inherent in positive rule making. Nevertheless, the GATT trade relations system, now encompassed in the 1995 World Trade Organization Agreement, proved an enduring idea and continues to be at the centre of the modern trade relations system. It was based on several key assumptions:

- trade policy should be non-discriminatory, as expressed in the principles of unconditional most-favoured-nation treatment and national treatment;

---

also John Croome: *The Results of the Uruguay Round: A Guide*, (Geneva: WTO, 1995).

[12] We explore the detail of what this entailed in "Post-Modern Trade Policy: Reflections on the Challenges to Multilateral Trade Negotiations After Seattle," *Journal of World Trade* 34:3 (June 2000).

- the primary regulator of trade should be the highly visible mechanism of the tariff, a mechanism that affects prices, rather than other mechanisms such as quantitative restrictions;
- tariffs and other barriers to trade should be progressively reduced so that the gains from trade can be realized and economic welfare increased;
- trade results largely from the activity of private entrepreneurs rather than governments;
- governments should be allowed to impose barriers against dumped or subsidized or otherwise politically intolerable levels of imports; and
- disputes between members should be resolved through a process of consultation and negotiation; retaliation and counter-retaliation should be avoided.

Over the years, the system became increasingly complex in order to compensate for the failure of the more ambitious International Trade Organization (ITO) to come into being. Despite these difficulties, the cumulative impact of the GATT—and now the WTO—was significant. Merchandise trade was liberalized, particularly among the advanced industrial economies; tariffs were cut; old-fashioned discriminatory quantitative restrictions were eliminated; and many potentially harmful practices were restrained by its rules[13].

---

[13] Its impact was much more limited in curbing the protectionist instincts of developing country governments. As a result, their participation in the benefits of international trade was much more limited, a reality that is only now beginning to be addressed by some developing countries through unilateral measures and bilateral arrangements. We explore the perverse impact of special and differential treatment for developing countries and their governments' approach to multilateral negotiations in "Special and Differential Treatment and the Doha 'Development' Round," *Journal of World Trade* 37:2 (April 2003). Gary Hufbauer sarcastically points out, "everyone 'knows' that trade ministers representing poor countries can't be asked to dismantle their barriers because ... well, because they like to use muddled infant industry arguments to confer favours on well-connected constituents." "Inconsistency between Diagnosis and Treatment," *Journal of International Economic Law*, 8:2 (June 2005), p. 293.

Mercantilist bargaining proved a politically successful way to harness the power of comparative advantage and its benefits can be clearly explained on the basis of conventional trade theories, particularly the Heckscher-Ohlin-Samuelson model. But these models rely on a stylized explanation of the mechanics of production and international trade that is increasingly far removed from current experience in the real economy. The way firms are organized to produce goods and services for either domestic consumption or export is of little moment to trade theory; industrial organization (IO) theory can be used to describe it, but few bridges were built to integrate the theories developed in these two branches of micro-economics until recently. Industrial organization theories explain the organization of production while international trade theories explain the exchange across national borders of the resulting goods and services. Traditional trade policy thinking and practice continues to rely on this division of labour and remains poorly informed on emerging IO theory.

The old trade policy assumed that international trade, investment, and other cross-border transactions took place largely between firms and individuals in one country and unrelated firms and individuals in another. It regarded the large, vertically integrated, horizontally diversified, managerially coordinated enterprise famously described by Alfred Chandler as typical[14]. With the exception of raw materials, machinery, and luxury products, experience suggested that goods and services were primarily consumed in the country in which they were produced by such firms. Goods and, to some extent, services were also produced for export but the successful penetration of a market often led to import-substituting investments and a return to the

---

[14] Alfred Chandler, *The Visible Hand: The Managerial Revolution in American Business* (Cambridge: Harvard University Press, 1977) and *Scale and Scope: The Dynamics of Industrial Capitalism* (Cambridge: Harvard University Press, 1990). Chandler's work was given important theoretical underpinnings in Oliver Williamson, *The Economic Institutions of Capitalism* (New York: The Free Press, 1985), relating the structure and operation of these firms to emerging theories about transaction costs, imperfect competition, and more.

dominant pattern of domestic production for local consumption. International transactions thus largely involved sales of primary goods, machinery, and luxury goods, exports of excess production, establishment of foreign affiliates through foreign direct investment, transfers of technology through affiliates and licensing arrangements, and the provision of various supporting services such as transportation, communication, and insurance. Nevertheless, in response to market liberalization, particularly among members of the Organization for Economic Cooperation and Development (OECD), exports steadily increased as a share of global output, rising from about 7 percent in 1950 to about 15 percent by the mid-1970s[15].

Lessons from Regional Negotiations

Mercantilist bargaining was not the basis upon which the members of the European Union negotiated the treaties that formed the basis for their increasingly integrationist project, nor was it the basis for concluding the Canada-United States Free Trade Agreement or the North American Free Trade Agreement. Each of these negotiations sought a much more thorough approach to rule-making and the resulting agreements reflect a detailed construction of rules to govern deepening integration among its members. They set out to cover the full range of cross-border transactions, from labour and capital to services and intellectual property rights. From the outset, these negotiations aimed at encouraging efficiency and integration, rather than simply the mercantilist goal of increased exports. In some ways, these re-

---

[15] The WTO calculates that the ratio of world trade in goods and services to output increased from 7 to 15 per cent over the period 1950 to 1974, and from 15 to 28 per cent between 1974 and 2004 , i.e., it has quadrupled since 1950. It grew most rapidly in the first two decades, slowed perceptibly during the 1970s and 1980s, and again grew rapidly since, consistent with, first, the impact of post-war recovery, and second, the impact of regional and global integration. See WTO, *International Trade Trends and Statistics* (Geneva: WTO, 1996 and 2006), accessed at wto.org. A recent *Economist* survey focuses on the extent to which the more recent surge has integrated developing country producers more fully into global production, trade, and investment patterns. "The new titans," *The Economist*, September 14, 2006.

gional negotiations reflected the full implementation of the logic of the old trade policy, but in other ways tried to come to grips with the demands of promoting and governing deeper integration.

Against a background of nearly half a century of depression and war, European governments in the 1950s embarked on an ambitious program of political cooperation and economic integration. Based on the deeply held conviction that countries that trade with each other and have an interest in each others' economic welfare are less likely to go to war or engage in destructive protectionist strategies, western European governments pursued top-down, policy-induced economic integration[16]. Over the course of the past five decades, the European integration movement has steadily expanded from the original six to now 27 member states, plus association arrangements with neighbours, potential members, and former colonies.

Simultaneous to the widening of the EC/EU, member governments steadily worked to deepen its impact. The 1957 Treaty of Rome committed members to implement the four freedoms: free movement of goods, services, capital, and people. Implementation of the free movement of goods was effected by removing intra-European tariff and non-tariff barriers and by adopting a Common External Tariff and a Common Agricultural Policy. This was accomplished by the original six members by 1968 and became a condition of entry for all subsequent members. The free movement of the other factors of production, however, proved a much more daunting challenge. In effect, it required a high degree of convergence in the regulatory regimes that are at the heart of the modern welfare state and that, either directly or indirectly, operate to segment national markets and frustrate integration. Over the past 20 years, through a process of both legislation and litigation, members of the EU succeeded

---

[16] Noted Jean Monnet, the father of European integration, in one of the most-quoted passages from his Memoirs "There will be no peace in Europe if States reconstitute themselves on a basis of national sovereignty. ... European States should form themselves into a federation or a 'European entity' which would make them a joint economic unit." *Memoirs* (Garden City, NJ: Doubleday, 1978).

115

in creating the rules and structures required to govern a continent-wide single market. As Michelle Egan points out, by closely emulating the US experience in creating an integrated, continent-wide market in the 19th century, the EU succeeded in erasing borders and creating a similarly integrated market in Europe. In the United States, the Constitution's Commerce clause and the courts proved critical, while the EU relied on the Treaty of Rome and the European Court of Justice[17].

The two North American Agreements, on the other hand, reflected an effort by first Canada and the United States, and then Mexico, to catch up to the reality of market-driven, cross-border integration by negotiating a framework of rules consonant with that reality. Rather than the push of government action, Canada-US integration has been driven largely by the pull of market forces: proximity, consumer choice, investment preference, and firm behaviour. Government policy has been largely responsive, motivated by efforts to resolve problems generated by market-driven integration. Rather than seeking deeper integration, governments only gradually accepted the need to facilitate it by addressing problems experienced by private traders and investors. The result is a much more piecemeal and less deliberate approach to rule-making and institution-building. Unlike in Europe, the governmental response in North America has been prompted by commercial and economic considerations and has been at pains to keep geopolitical and security considerations at arm's length in forging new rules and arrangements to address deepening economic integration.

---

[17] See Michelle P. Egan, *Constructing a European Market: Standards, Regulation, and Governance* (New York: Oxford University Press, 2001). The American parallel is instructive, particularly in light of Canada's own more difficult progress in effecting a more integrated market through the Internal Trade Agreement. For Canadians, perhaps the most instructive lesson from Europe does not lie in efforts to create first a customs union and then a single market, but in the efforts of the Outer Seven—the members of the European Free Trade Association (EFTA)—first to maintain their separate markets, then to negotiate individual free-trade agreements followed by the short-lived European Economic Area (EEA) arrangement and finally full participation in the single market for most EFTA members. The logic of the market drove decisions about its governance.

British economist David Henderson defines integration "as a tendency for the economic significance of political boundaries to diminish."[18] It takes place when one or more formerly separated markets combine to form a single market, leading to increased flows of cross-border trade of not only final consumer products but also intermediate inputs and raw materials, as firms reorganize their activities around regional markets for both inputs and outputs. Integration enables producers and consumers to benefit more fully from their relative strengths and to respond more efficiently to changing economic conditions. It may open new territories for an economy's output, strengthen exploitation of economies of scale, increase access to potentially cheaper suppliers of inputs, and create new opportunities for foreign direct investment, as firms restructure the vertical and horizontal arrangements of their enterprises. Greater competition is also likely to make goods and services more affordable, thereby expanding consumer purchasing power, and add to consumer choice.

Whether driven by the push of government policy or the pull of market forces, deepening integration between two or more countries disposes them to create common policies to regulate the production and distribution of goods and services and a joint approach to external trade and investment. Canada and the United States, while nominally committed to no more than a free-trade area, have in reality already implemented aspects of a customs union and even of a common market. Based on broadly shared goals and perspectives and common needs, the two governments have developed a dense framework of formal and informal networks and relationships that ensures a high degree of convergence in the design and implementation of a wide range of rules and regulations.

Both the European treaties and the North American agreements are built on an architecture of positive rules rather than negative prescriptions, and neither was the product of mercantil-

    [18] David Henderson, "Putting 'Trade Blocs' into Perspective," in Vincent Cable and David Henderson, eds., *Trade Blocs? The Future of Regional Integration* (London: Royal Institute of International Affairs, 1994), pp. 179-80.

ist bargaining. Both are much more ambitious in their coverage. Both anticipate the new trade policy and both reflect the much more intensive nature of exchange within the territory covered by the rules. In the case of Europe, the treaties helped to forge more integrative business strategies. In the case of North American, the agreements reflected the extent to which business was pursuing more integrative strategies. In both instances, the architecture and bargaining strategy of the old trade policy proved ill-suited.

The effective market today is global and its reach is increasingly reflected in the organization of production. The United States in the 19th century and the EU, and to a lesser extent, Canada in the closing years of the 20th century, saw a need to forge rules and governance structures consonant with the emergence of larger markets and more widely integrated production strategies. The focus of these efforts was divergent regulatory regimes that artificially segmented markets and frustrated achieving the benefits of wider markets and more efficient production structures. Today, both markets and production have gone global, while governance remains largely national in scope and reach.

The New Industrial Context: Fragmentation and Integration

Over the past few decades, traditional international exchange has gradually begun to give way to a much more integrated kind, with more and more transborder transactions taking place within firms, among related parties, or within integrated networks[19]. Many more goods traded internationally today are parts and components for assembly into end-products closer to the point of final consumption. Production is being geared to a much wider market, the range of goods and services that are

---

[19] The UNCTAD Division on Transnational Corporations and Investment reports that by 2005, some 77,000 firms qualified as multinational in their activities, each accounting for an average of ten separate foreign affiliates. World-wide sales by foreign affiliates had reached US$22.2 trillion in 2005, nearly double world-wide exports of goods and services at US$12.6 trillion. See World Investment Report (Geneva: UNCTAD, 2006).

exchanged internationally has widened considerably, and capital and technology move between nations not only to promote import-substituting, but also export-oriented production. Global competition, scientific and technological breakthroughs, as well as consumer sophistication are shortening the product cycle and placing a premium on quality, manufacturing fluidity, and in-innovation. International exchange now involves a much more complex and sophisticated range of economic transactions and is as likely to involve dealings among related than unrelated parties. The vertically integrated firms of the early postwar years have given way to much more flexible, horizontally organized enterprises[20]. Production is steadily being re-organized on a global basis and the nature of extra-national economic transactions reflects this change. In the words of the University of Manchester's Peter Dicken, the global economy has been transformed into "a highly complex, kaleidoscopic structure involving the fragmentation of many production processes, and their geographical relocation on a global scale in ways which slice through national boundaries."[21]

The literature suggests that there were three basic catalysts to the acceleration of globalization: the steady liberalization of trade and investment among industrialized countries after the Second World War, the more recent rapid industrialization of the third world, and the impact of technological breakthroughs that have brought down the costs of transportation and communication. The impact of these three factors has been mutually reinforcing and cumulative. Duke University sociologist Gary Gereffi insists, however, that "of far greater significance are several novel features in the nature of international trade that do not have counterparts in previous eras. ... The three new aspects

---

[20] See Naomi R. Lamoreaux, Daniel M.G. Raff, and Peter Temin, "Beyond Markets and Hierarchies: Toward a New Synthesis of American Business History," *American Historical Review* 108 (April 2003), pp. 404-33. Much of the article is a critical analysis of the lessons of business history since Chandler wrote *The Visible Hand*.

[21] Peter Dicken, *Global Shift: Reshaping the Global Economic Map in the 21st Century* (4th edition, London: Sage, 2003), p. 9.

of modern world trade relevant here are (1) the rise of intra-industry and intra-product trade in intermediate inputs; (2) the ability of producers to 'slice up the value chain,' in Krugman's phrase, by breaking a production process into many geographically separated steps; and (3) the emergence of a global production networks framework that highlights how these shifts have altered governance structures and the distribution of gains in the global economy."[22]

The fragmentation of production through a process of outsourcing and subsequent re-bundling within large and technologically sophisticated supplier networks has become increasingly prevalent in industries from food processing, aviation, and motor vehicles to apparel, electronics, and house-hold products[23]. Both value-chain fragmentation and the sophistication of the firms that make up the fragments have made it easier to relocate specific nodes of production and to take advantage of a range of distant factors, from low-cost labour and specialized skills to access to critical inputs and public policy considerations. As the production of manufactured goods becomes ever more disaggregated, varied, and sophisticated, the cost of developing and manufacturing new products has increased exponentially. More and more, the costs are concentrated in developing the product—both the product and the most cost-effective process by which to manufacture it—rather than in production, devaluing the labour content in many products and increasing the risk in producing it. MIT geographer Tim Sturgeon points out: "In both manufacturing and service industries, ... many

---

[22] Gary Gereffi, "The Global Economy: Organization, Governance, and Development," chapter 8 in Neil J. Smelser and Richard Swedberg, eds., *The Handbook of Economic Sociology*, 2nd edition (Princeton, NJ: Princeton University Press and Russell Sage Foundation, 2005), p. 166.

[23] Sven W. Arndt and Henryk Kierzkowski note "fragmentation is not a new phenomenon; nor is out-sourcing. ... In the modern era, however, both have acquired international dimension and complexity and probably represent one of the most important distinguishing features of contemporary globalization." "Introduction," in Arndt and Kierzkowski, eds., *Fragmentation: New Production Patterns in the World Economy* (Oxford: Oxford University Press, 2001), p. 2.

companies have been shifting specialized activities out-of-house to an increasingly competent set of suppliers, contract manufacturers, and intermediaries. ... While offshore assembly was initially done by the subsidiaries of multinational firms, growing capabilities in the supply-base led to the emergence of independent and highly sophisticated developing country suppliers in places like Taiwan, Korea, and Singapore as well as a cadre of huge 'global suppliers' headquartered mainly in developed countries but with extensive worldwide operations." [24]

Li & Fung of Hong Kong is a prime example of a modern specialist in managing the economics of fragmentation and agglomeration. It has access to a network of about 7,500 contract suppliers all over Asia employing as many as 1.5 million workers, providing a critical mediating service bringing brand-name firms together with highly efficient contract manufacturers. Li & Fung takes orders from companies all over the world to "make things" for them, from ballpoint pens and golf clubs to computers and televisions. It in turn finds the right contract manufacturer and organizes the logistics to supply the ordered "thing" to the customer, based on the customer's specifications. Many contract suppliers maintain offices in Hong Kong to liaise with Li & Fung and provide it with product development and engineering information [25].

Firms like Li & Fung are key to understanding the increasing role of China in the value-added chain: the location of choice for "making things." India, on the other hand, has become the favoured place for service inputs. An increasing range

[24] Timothy Sturgeon, *Measuring the Offshoring of Service Work and its Impact on the United States*, accessed at http://web.mit.edu/ipc/owg2005/summary.html.

[25] See Victoria Curzon Price, "Some Causes and Consequences of Fragmentation," in Arndt and Kierzkowski, Fragmentation: *New Production Patterns in the World Economy* (Oxford: Oxford University Press, 2001), p. 96. As its website boasts, "Li & Fung is today one of the premier global consumer-products export trading corporations managing the supply chain for high-volume, time-sensitive consumer goods. Our mission is to deliver the right product at the right price at the right time." www.lifung.com. See also "Manufacturing Survey," *The Economist*, June 20, 1998.

of service inputs are being sourced in India, taking advantage of its wealth of IT professionals and English-speaking, well-educated office workers. Its contribution began with low-value-added activities, such as back-office transactions and call centers, but steadily expanded to include software programming, engineering, design, accounting, legal and medical advice, and a broad array of other professional services. "The computerization of work, the advent of the Internet, the standardization and automation of a range of business processes, and the availability of very low-cost, high-speed data networks have made it easier for firms in advanced economies like the United States to connect to the capabilities that exist in developing countries like China, Ireland, Australia, Canada, India, and the Philippines, and this combination of factors appears to be causing a wide range of 'knowledge work' to become more footloose."[26]

Fragmentation and integration combine to increase the extent and intensity of international transactions, allowing slices of the production process to be moved to the best possible location. The speed and efficiency with which these slices can be integrated clearly have a bearing on the optimal degree of fragmentation. Fragmentation thus allows firms to specialize to a much greater degree and reap greater advantages from economies of scale and scope. Gereffi points out that "today, we live in a world in which deep integration, organized primarily by transnational corporations (TNCs), is pervasive and involves the production of goods and services in cross-border value-adding activities that redefine the kind of production processes contained within national boundaries. … While the postwar international economic order was defined and legitimized by the United States and the other core powers that supported it in terms of the ideology of free trade, it was the way in which transnational corporations linked the production of goods and services in cross-border, value-adding networks that made the

---

[26] Sturgeon, *Measuring the Offshoring of Service Work and its Impact on the United States*, op. cit.

global economy in the last half of the twentieth century qualitatively distinct from what preceded it." [27]

Systematic data on the extent of this integration is difficult to find, in part because official statistics cannot capture the full value of cross-border service links or the input of services provided through proprietary and other networks, e.g., design, engineering, and marketing, whether done in-house, outsourced locally, or outsourced internationally. Statistical agencies have no way of counting the value of Italian design and German engineering in a toilet ultimately manufactured in Mexico and imported into Canada through a US distribution network. They count the computer on which this is being written as a Chinese import, rather than as the fruit of the design, engineering, and marketing input of the brains at Apple's Cupertino, California campus[28]. The data collected by statistical agencies thus does not always accurately capture the origin of value-added in a complex multi-country production process[29].

From a policy perspective, governments are particularly interested in the intersection of firm-specific value and location-

---

[27] Gereffi, "The Global Economy: Organization, Governance, and Development," op. cit., at pp. 163-64.

[28] As Business Week points out, the statistical wizards at national statistical agencies "have no way of tracking the billions of dollars companies spend each year on innovation, product design, brand-building, employee-training, or any of the other intangible investments required to compete in today's global economy." Michael Mandel, et al., "Why the Economy is a Lot Stronger than You Think," *Business Week*, February 13, 2006. The increasing proportion of services in national accounts reflects not only a more prosperous economy with an increasing desire for services rather than goods, but also the disaggregation of production into increasingly smaller slices, many of which are counted as the production of services rather than as part of the goods in which they are embedded, as well as the service links required to bring the spatially separated slices together.

[29] Alexander Yeats, by analyzing data for selected industries and extrapolating the results more widely, estimates that a third or more of world trade is made up of parts and components. "Just How Big is Global Production Sharing?" in Arndt and Kierzkowski, eds., Fragmentation: New Production Patterns in the World Economy (Oxford: Oxford University Press, 2001), pp. 108-143.

specific value. Firms are now less constrained in their choices of geographic location by distance and policy, and seek to enhance value by dispersing their value-added activities spatially. Governments, in the interest of attracting value-added activity to their location-specific jurisdictions, now compete in promoting policy settings that are congenial to increasingly mobile slices of production by removing barriers and providing positive incentives. In this quest, they are learning that the trade agreements of the past may have been critical to providing the framework of rules that initially promoted fragmentation and integration, but are no longer sufficient.

Canada-US Cross-Border Integration

The integration that increasingly characterizes the global economy has a longer history at a bilateral Canada-US level. In an earlier era, proximity disposed Canadians to develop a trade and investment dependence on the US market and US capital. The exploitation of Canada's storehouse of raw materials and the establishment of miniature-replica branch plants both developed with heavy doses of US capital. Today, Canada exhibits a high level of both production and consumption dependence on the United States[30]. Its earliest modern manifestation involved the automotive industry. A unique set of circumstances at the time, captured in the Autopact, encouraged the development of integrated, cross-border production[31]. Much of what is now commonplace was pioneered in the auto sector: in-house fragmentation and outsourcing on a continental rather than regional or national basis, followed by out-of-house cross-border fragmentation. The successful introduction of lean, just-in-time production techniques, pioneered in Japan and introduced in North

---

[30] For a description of the geography of production in Canada, see Iain Wallace, *A Geography of the Canadian Economy* (Toronto: Oxford University Press, 2002).

[31] On the origins and evolution of the Autopact, see Michael Hart, *A Trading Nation: Canadian Trade Policy from Colonialism to Globalization* (Vancouver: UBC Press, 2002), pp. 240-47.

America in the 1980s, further accelerated this fragmentation process[32].

Since the implementation of the CUSFTA and NAFTA, fragmentation has become commonplace throughout North American industry, involving both manufacturing and service industries[33]. High levels of both two-way intra-industry trade and foreign direct investment indicate continued cross-border integration and rationalization of production between Canada and the United States, as well as a deepening interdependence of manufacturing industries. Canada is the second-leading destination for US foreign direct investment while the United States is the prime destination of Canadian FDI. Proximity of the US and Canadian industrial heartlands, well-developed infrastructures, and transparent legal systems all contribute to the highly integrated nature of the two economies. In turn, this integration contributes to a high level of trade as each country is the other's largest foreign market and leading supplier of imported goods[34].

Discussion of Canadian international economic patterns often focuses on trade in goods and emphasizes exports. A more realistic picture emerges, however, by looking at imports and exports of both goods and services, at inflows and outflows of

[32] See James P. Womack, Daniel T. Jones, and Daniel Roos, *The Machine that Changed the World* (New York: Macmillan, 1990).

[33] For example, the Canadian and US agri-food sectors have become much more integrated. Canadian and US fresh fruit and vegetable consumption is growing and increasingly involves two-way trade, with the direction dependent on the season. Cross-border production by Canadian and US multinational food companies has steadily risen, lowering production costs and giving consumers access to a wider variety of products. See Steven, Zahniser, ed., *NAFTA at 11: The Growing Integration of North American Agriculture* (Washington, USDA Economic Research Service Outlook Report No. WRS0502, February 2005), accessed at www.ers.usda.gov/Publications/WRS0502/ .

[34] See U.S. International Trade Commission, *Production Sharing: Use of U.S. Components and Materials in Foreign Assembly Operations, 1995-1998*, Publication 3265 (Washington, 1999), for a detailed assessment of the extent and nature of cross-border production sharing in the motor vehicle, aircraft, rail locomotives and rolling stock, computer hardware, semiconductor, and telephone equipment industries.

investment capital, at sales by foreign affiliates, and at exports of goods as a share of domestic shipments. As Howard Lewis and David Richardson point out, "it is becoming increasingly meaningless, if not outright impossible, to think of trade as something separate from cross-border investment, or of exporting as something separate from importing products and innovative ideas. All are tied together in the extended family of global commitment."[35] As such, Canada's involvement in the global economy is much more diversified and the full importance of international exchange becomes clearer. It also makes clear why, as the US economy moves further up the value chain, so does the Canadian economy, increasing trade opportunities for foreign exporters to North American markets and investment opportunities in overseas economies.

Preliminary figures for 2006[36] indicate that Canadian firms and individuals produced $524 billion in goods and services for export and imported $486 billion in goods and services from all sources. The United States remained by far the most important destination for Canadian merchandise exports at $362 billion (78.9 percent), and the principal supplier of Canadian merchandise imports at $265 billion (65.5 percent). The increasing value of the Canadian dollar relative to the US dollar as a result of rising resource prices appears to have had a larger impact on diversifying merchandise imports rather than exports. China has become the second largest supplier to the Canadian market, while the EU remains the second largest destination for Canadian exports.

On the investment front, Canadian direct investment abroad, with a cumulative book value of $465.1 billion in 2005, continued to outpace the cumulative book value of foreign direct investment in Canada at $415.6 billion[37]. The value of Ca-

---

[35] Howard Lewis III and J. David Richardson, *Why Global Commitment Really Matters* (Washington: Institute for International Economics, 2001), p. 11.

[36] Figures from Foreign Affairs and International Trade Canada, http://www.international.gc.ca/eet/balance-payments-en.asp.

[37] Figures from Foreign Affairs and International Trade Canada, http://www.international.gc.ca/eet/foreign-statements-en.asp.

nadian direct investment abroad exceeded the value of foreign direct holdings in Canada for the first time in 1997 and has remained that way ever since. As a result of increasing cross-border investment, Canadian firms are increasing their presence in the United States through sales by affiliates, particularly in the services sector, just as US-owned affiliates continue to have an important place in the Canadian market place. Analysis of cross-border investment patterns indicates that much of it is trade-enhancing as Canadian and US firms strengthen their position in supply chains and distribution networks while overseas investment is geared more to substituting for trade. McCain's, for example, invests in Europe to process locally sourced inputs while it invests in the United States to enhance its ability to distribute product from its Canadian operations. Canadian firms have become increasingly involved in cross-border mergers and acquisitions, the principal vehicle for FDI flows and for seizing the advantages of deepening integrative trade. From the beginning of 2003 through the first quarter of 2006, Canadian firms acquired more than 1,000 foreign firms, while foreign firms acquired 373 Canadian firms.[38]

In 2003, the latest year for which such data are available, about 43,310 Canadian-based firms were engaged in exporting. Of these, 27,747 (64.2 percent) exported to the United States alone, 5,802 (13.4) percent) exported only to non-US destinations, and 9,671 (22.4 percent) to both. Fewer than 4 percent of these firms, accounted for more than 80 percent of total exports, and 12 percent accounted for 93.5 percent of Canada's total merchandise exports. The profile of successful exporting firms is also highly skewed toward foreign-controlled firms. Of the 44,469 firms which registered as exporters in 2002, only 3,597 (9 percent) were foreign-controlled, but they were responsible for nearly half of merchandise exports[39].

---

[38] Robert Yalden, "Gobble Means Grow," *National Post,* June 28, 2006, p. FP23.

[39] Statistics Canada, *A Profile of Canadian Exporters 1993-2003,* Catalogue No. 65-506-XIE, and Craig Byrd, Foreign Control of Canada's Merchandise Exports, 2002 (Statistics Canada Catalogue No. 65-507-MIE, No. 4)

Similar data are not available on the number of Canadian-controlled exporting establishments with assets or affiliates in foreign markets, particularly the United States, or of the ownership and investment profile of importing establishments, but the data on the dominant position of foreign-controlled exporters suggest that firms active in international markets as traders are also active in those markets as investors. As Lewis and Richardson discovered, internationally engaged firms are more productive and innovative, pay higher wages, and are more profitable, not just in the United States, but in most other economies[40].

Both cross-border and global supply chains today depend critically on relationships that extend well beyond arm's-length transactions between customer and supplier. As US business economist Stephen Blank notes "Ottawa and Washington talk about the world's largest bilateral trading relationship. But we really don't trade with each other, not in the classic sense of one independent company sending finished goods to another. Instead we make stuff together; … [we] share integrated energy markets; dip into the same capital markets; service the same customers with an array of financial services; use the same roads and railroads to transport jointly made products to market; fly on the same integrated airline networks; and increasingly meet the same or similar standards of professional practice."[41]

Philip Cross and his colleagues at Statistics Canada have done extensive work trying to understand the changing nature of cross-border trade, production, and investment patterns. They have calculated that the import content of Canadian exports has risen steadily over the past two decades. It was 25.5 percent in 1987 and peaked at 33.5 percent in 1998. The rapid rise in trade in the 1990s was in large part the result of rationalization, with imported components replacing domestic components, with the final product exported to a broader market base. More significant than the rise in exports as a share of GDP was the rise in

[40] Lewis and Richardson, *Why Global Commitment Really Matters.* Op. cit.

[41] Stephen Blank, "It is Time for Canada to Think Carefully about North America." *Embassy.* September 7, 2005.

value-added content of exports in GDP, which rose from 21.4 percent in 1987 to reach 28.8 percent in 1999[42]. The recent increase in the value of sales of energy products has had a dampening effect on a further rise in the import content of Canadian exports, but not on their Canadian value-added content.

The extent of the import content of Canadian exports varies considerably from industry to industry. Not surprisingly the auto sector, benefiting from high levels of cross-border ownership since its inception and, since 1965, the impact of the Canada-US Autopact, is by far the most integrated, with the import content of Canadian exports exceeding 50 percent, followed by machinery, equipment, and electronics at over 40 percent, and textiles, other manufacturing, metals, oil refining, and chemicals exceeding 30 percent. Even food, forestry products, and agriculture exceeded 10 percent[43]. Economist Glen Hodgson concludes that "trade has evolved beyond the traditional exporting and importing of goods, and has entered the next generation of trade—integrative trade. Integrative trade is driven by foreign investment and places greater weight on elements like the integration of imports into exports, trade in services and sales from foreign affiliates established through foreign investment."[44] Nowhere has this process of integration been more pronounced than between Canada and the United States.

The rise in Canada's export dependence on the United States in the 1990s was, therefore, to some extent overstated. As a result of double counting, net exports to the United States as a share of GDP only rose to reach 24 percent by 1999, as com-

---

[42] See P. Cross and G. Cameron, "The Importance of Exports to GDP and Jobs," *Canadian Economic Observer*, 12:11 (November 1999) and Cross, "Cyclical Implications of the Rising Import Content in Exports," *Canadian Economic Observer*, 15:12 (December 2002).

[43] See Figure 2 in Cross "Cyclical Implications of the Rising Import Content in Exports."

[44] Glen Hodgson, "Trade in Evolution: the Emergence of Integrative Trade," *EDC Economics*, March 2004, p. 5; http://www.edc.ca/english/docs/030104_integrative_e.pdf. See also Hodgson, "Integrative Trade and the Canadian Experience," *EDC Economics*, May 2004, http://www.edc.ca/english/docs/Canadian_Benefits_050104_e.pdf.

pared to the 36 percent suggested by the gross numbers[45]. In other ways, however, these numbers may understate the degree of interconnectedness. Canadian merchandise exports as a share of domestic shipments have steadily increased over the past thirty years as Canadian firms have become much more engaged in cross-border and international supply chains. In the 1970s, the value of Canadian merchandise exports was equivalent to about two-thirds of the value-added in the production of goods in Canada. In the opening years of the 21st century, merchandise exports now represent about 125 percent of the value added in the goods-producing sectors of the Canadian economy, indicating nearly a doubling in the export intensity of production in Canada over the past thirty years, as well as the increasing role of imported components in that production[46]. Additionally, as a result of increasing cross-border investment, Canadian firms are increasing their presence in the United States through sales by affiliates, particularly in the services sector, just as US-owned affiliates continue to have an important place in the Canadian market place. Canadian-owned affiliates in the United States rang up $192 billion in sales in 2003[47], roughly 60 percent of the value of Canadian-based exports. In the other direction, US-owned affiliates in Canada reported 2003 sales of $560 billion[48].

Deepening integration has allowed Canadian industry to become more specialized and has contributed importantly to the growth of value-added sectors. As Industry Canada economist Surendra Gera and his colleagues explain, "while in the past domestic demand was the dominant factor influencing the

---

[45]   Cross and Cameron, "The Importance of Exports to GDP and Jobs."

[46]   Values derived from Statistics Canada, CANSIM Tables 376-0007, 379-0022, 379-0024, and 380-0012 reporting Gross Domestic Product and Balance of Payments data.

[47]   Statistics Canada, "Foreign Affiliate Trade Statistics." *The Daily*, May 25, 2005, at http://www.statcan.ca/Daily/English/050525/d050525f.htm.

[48]   US Bureau of Economic Analysis, *U.S. Direct Investment Abroad: Operations of U.S. Parent Companies and Their Foreign Affiliates: Comprehensive financial and operating data,* http://www.bea.gov/bea/ai/iidguide.htm#link12b; US dollar data converted to C$ at Bank of Canada average rate for 2003 of 1.4015.

growth of industries, trade is now becoming much more impor-
tant. High-knowledge industries in the tradable sector seem to
have benefited the most from export performance; low-
knowledge industries have seen their relative decline hastened
by import competition."[49]

The changing intensity and composition of bilateral trade
have contributed significantly to making Canadians better off
both as consumers and as producers. Canadians employed in
export-oriented sectors have consistently been better educated
and better paid than the national average. As University of To-
ronto economists Peter Dungan and Steve Murphy report,
"Canada is replacing low-productivity employment with high-
productivity employment through expanded international trade,
and is thereby made better off."[50] Similarly, greater access to
internationally competitive goods and services allows Canadians
to stretch their earnings further. As Cross and his colleagues
point out: "The importance of trade to the economy does not
come from an excess of exports over imports: rather, it is from
the productivity gains that accrue with increased specializa-
tion."[51] Specialization, in turn, increases, as markets expand in
response to the increased openness fostered by trade agreements.

Emerging Cross-Border Trade Policy

The relationship with the United States is thus the indispensable
foundation of any Canadian policy to maximize benefits from
engagement in the global economy. With no other country does
Canada have as intense a relationship embracing virtually the
whole range of public policy, security, economic development,

[49] "The Knowledge-Based Economy: Shifts in Industrial Output," *In-
dustry Canada Working Papers,* Number 15, January 1997, accessed at
strategis.ic.gc.ca.

[50] "The Changing Industry and Skill Mix of Canada's International
Trade," Industry Canada Research Publications Program, Perspectives on
North American Free Trade Series, Paper Number 4, April 1999, available at
http: strategis.ic.gc.ca epic site eas-aes.nsf vwapj P4-ang.PDF SFILE P4-ang.PDF

[51] Cross and Cameron, "The Importance of Exports to GDP and Jobs,"
p. 3.3.

and human contact. For a growing number of Canadians, the time has come to achieve a seamless border with our neighbour, embraced within a new agreement implementing rules, procedures, and institutions consonant with the reality of ever-deepening, mutually beneficial cross-border integration.

While some might express horror at any policy explicitly aimed at helping integration, there are no voices calling for the alternative—creating barriers to integration, establishing distance from the United States, and seeking other partnerships to replace this vital relationship. Who would support imposing barriers to the millions of visits that Canadians make to the United States each year, to the annual pilgrimage of several million snowbirds to Florida, Arizona, California, and other sunnier climes, to the thousands of trucks that cross the border every day, to the billions of cross-border phone calls, to the dozens of US TV channels beamed into Canadian homes, to the millions of US books, movies, CDs, and magazines delivered to Canadian homes every year? Who objects to building more cross-border roads, bridges, pipelines, and electrical grids? At a practical level, therefore, the default position is implicitly a gradual deepening of integration, whether stated in those terms or not.

The need for an integrated approach to managing the Canada-US relationship and for the focus upon border management is echoed by the Canadian business community. The Canadian Chamber of Commerce has called for the government to put in place the proper machinery to ensure the cross-border flow of goods and people. The Canadian Association of Manufacturers and Exporters similarly gives priority to improving border efficiency, eliminating border infrastructure bottlenecks, and reducing regulatory barriers to trade. The Canadian Council of Chief Executives argues the need to move beyond border management and reinvent the concept of North American Borders. It, moreover, places the border challenge in the context of reinvigorating the Canada-US defence and security relationship[52].

---

[52] See Chamber.ca for proceedings of its 2003 conference and the resolution calling for improved customs procedures. The position of the manufacturers is contained in a joint letter from it and its sister US group,

Canadians and Americans are now trying to come to grips with the fact that their extensive network of trade arrangements has worked exceedingly well, but that they no longer address directly the needs related to the circumstances in which the two countries find themselves. More profoundly, there is a need for new instruments. It is evident that neither the WTO as a body of rules nor the Doha Round of trade negotiations provide a way to address the emerging issues in the relationship in a timely or effective way. The issues are peculiar to the relationship, do not lend themselves to multilateral solutions, and, given the determination of developing countries to shape multilateral negotiating agendas, have no possibility of finding a place even if Canada and the United States were inclined to go down this road.

While a bilateral context offers the chance to shape the agenda and focus uniquely on Canada-US issues, it is far from certain, however, that a conventional new trade agreement, for example, a customs union,[53] would yield results significant enough to justify the investment of political capital necessary to

the National Association of Manufacturers, to the Prime Minister and President in April 2004, available at cme-mec.ca. See ceocouncil.ca for the Council's publication "New Frontiers: Building a 21st Century Canada US Partnership in North America."

[53] The past few years have witnessed a growing interest in what would be involved in negotiating either a bilateral or trilateral customs union involving Canada, the United States, and perhaps Mexico. We prepared an overview for the Policy Research Initiative, "Policy Implications of a Canada-US Customs Union," PRI North American Linkages Project, February 2005. See also Danielle Goldfarb, "The Road to a Canada-US Customs Union: Step by Step or in a Single Bound," *CD Howe Institute Commentary*, No. 184 (Toronto, 2003); Rolf Mirus and Nataliya Rylska, "Should NAFTA Become a Customs Union?" in Edward D. Chambers and Peter H. Smith, eds., *NAFTA in the New Millennium* (Edmonton: University of Alberta Press, 2002)); Axel Huelsemeyer, "Toward Deeper North American Integration: A Customs Union?" *Canadian American Public Policy*, No. 59 (Orono, ME: University of Maine, October 2004, and Hart "A New Accommodation with the United State: The Trade and Economic Dimension," in Thomas J. Courchene, Donald J. Savoie, and Daniel Schwanen, eds., *Art of the State II: Thinking North America: Prospects and Pathways*, No. 2 (Montreal: Institute for Research on Public Policy, March 2004).

achieve success. A customs union would build on the existing contours of the NAFTA, but extend it to include:

- a single customs territory, allowing for the free circulation of all goods within the customs union.
- a common external tariff (CET) created from merging the two tariff schedules and harmonizing on the lowest current rate of either country, with flexibility to maintain separate external rates over a transition period for a small list of products.
- provisions for sensitive sectors such as clothing (likely transition periods), agriculture (leaving the hardest issues to WTO resolution), culture (similarly leaving aside the principal issues not already covered by the WTO), and recognition of the importance of the energy sector;
- reconciling the differences in the two countries' current free-trade and preferential-trade arrangements with other countries respecting product coverage, rules of origin, and possible future free trade partners.
- a common approach to trade remedy laws that would imply a single regime for third countries, and recognizing that Canada-US cases are rare, concern largely resource products, notably softwood lumber, and point to the need to pay attention to resource management issues.
- a common external trade policy with respect to multilateral and regional negotiations (which, given the degree of policy convergence, should not prove difficult) and more generally for the conduct of trade relations with third countries; with respect to trade sanctions for foreign policy reasons, it should be possible to maintain policy flexibility and devise practical instruments for this purpose without compromising the integrity of the union;
- dispute settlement within the customs union based upon the current NAFTA provisions with, as necessary, some major updating; and
- institutions for administering the union, resolving disputes, and facilitating dynamic, joint policy development to govern accelerating economic integration, drawing upon existing models such as the International Joint Commission.

Each of these would pose technically complex challenges but none would raise insurmountable barriers. The major policy issues inherent in implementing a customs union have essentially been resolved by a process of policy convergence and multilateral and bilateral trade agreements. This conventional approach would, in effect, clean up the leftovers from earlier negotiations but fail to address emerging issues flowing from deepening integration. The advantage of convention is the high level of comfort that Canadians have with the tried and true. The constraint is that the returns from such an approach promise to be minimal. Conventional trade policy provided an appropriate answer to the policy challenges of the past. It provides at best a partial answer to the challenges and problems that flow from the deep integration that now characterizes exchange between Canada and the United States. At its most basic, it would be limited to the exchange of goods, and not address issues related to the exchange of services, capital, or intellectual property rights, and not cover the movement of people.

Canadian trade policy, which in the past has provided answers to the conundrums of the relationship with the United States, is stuck in neutral. During the 2006 election, there was no public discussion of the deep integration that already exists and the broad patterns of cooperation and policy coordination with the United States that have emerged beneath the radar screens of political approval or formal agreement. The logic of Canada's economic interest makes a compelling argument to focus energies upon reinventing the US relationship to conform to modern realities. Yet, time stands still for no policy and unless answers are found to the quandaries of economic integration and the implications that flow from it, Canadians will begin to pay increasingly heavy economic costs from policy paralysis.

Reaping the full benefits of deepening cross-border economic integration will require that Canada and the United States address three fundamental, and interrelated, issues: reducing the impact of the border, accelerating and directing the pace of regulatory convergence, and building the necessary institutional capacity to implement the results of meeting the first two challenges. Each of these will prove difficult and solving the prob-

lems associated with either of the first two will prove illusory without addressing the other two.

Border Administration

The first challenge is to address the increasing dysfunctional border administration[54]. The intensity of the cross-border relationship is apparent from the 36,000 trucks and 400,000 people who cross the border every day. Nevertheless, even after 15 years of "free" trade, the Canada-United States border continues to bristle with uniformed and armed officers determined to ensure that commerce and interaction between Canadians and Americans complies with an astonishing array of prohibitions, restrictions, and regulations. The list of rules and regulations for which the border remains a convenient, and even primary, enforcement vehicle has grown, rather than diminished, since the implementation of free trade, particular in response to the new security realities created by 9/11. Administering the physical border imposes high costs on the two governments, on firms and individuals who use the border frequently to conduct their affairs in the integrated North American economy, and on the two economies in terms of opportunities missed to reap the full benefits of deep integration[55]. The result is a not an integrated,

---

[54] For more detail, see William B. P. Robson and Danielle Goldfarb, "Risky Business: U.S. Border Security and the Threat to Canadian Exports." *C.D. Howe Institute Commentary,* No. 177 (Toronto, March 2005).

[55] Firms in considering fragmentation strategies consider tariff and other border barriers, wage differentials, transport and other integration costs, input costs, government policies, and more. For example, a Canadian firm supplying customers throughout North America, designs and engineers its products in Canada, sources high-end components in the US and Japan, lower-end components in Malaysia and Korea, and contracts for final assembly in China. Product is then shipped to North America for distribution to customers in both Canada and the United States, who rely on local distributors for after-market service. Location of the central distribution facility will in part be determined by the MFN tariff and rules of origin that will apply, and will likely take place in the United States, even if a Canadian location has clear logistic and other cost advantages, because the bulk of product can be distributed on the basis of one customs clearance rather than two. The result is not important to US policy makers but is critical to Canadians.

single North American market, but two markets with many cross-border ties that remain hostage to the efficiency and reliability of customs clearance, an issue of greater importance to Canadian-based than US-based firms.

In addition to routine customs and immigration activities, both Canada and the United States use border controls to interdict illegal immigration, drugs, terrorism, and other criminal activities. Experience, however, suggests that the cost of border administration to pursue these goals is out of proportion to the results[56]. The border is simply too long and too porous to prevent determined cross-border criminal activity. Devoting even more resources at border checkpoints along the bilateral border seems unlikely to achieve additional results absent extraordinary further investments in human and physical infrastructure. Increasing resources to such an extent, however, risks causing considerable collateral damage to economic interests in an effort to find solutions to a problem that can be handled more effectively and efficiently through other initiatives.

To that end, ways need to be found to reduce the impact of the border. The two governments could, for example, strengthen institutional contacts, enhance cooperation, and share information on matters small and large. They could explore further investments in intelligence gathering and gradually focus ever larger parts of that effort at initial entries into North America. They could also make greater investments in infrastructure and in technology (both at ports-of-entry and the corridors leading to such ports). Both types of investments are critical components of any comprehensive effort at improving the management of the border and reducing its commercial impact. Such investments need not proceed on the basis of current inspection methodologies, but could rely much more on risk assessments and random inspections[57]. They could also focus more on tar-

---

[56] This, of course, is not a view shared by all; some hold that no cost is too large to protect the country from illegal drugs, immigrants, and other criminal activity. This perspective was well represented by Tim Naumetz, "Summer border policy: take operational risks," *National Post*, 22 July 2003 A2.

[57] To combat terrorism and other illegal activity, for example, Canada and the United States need rapid and timely exchanges of information on

geting resources toward pre-clearance programs for goods, vehicles, and people. Finally, the two governments could enhance discussions about increasing the level of convergence in US and Canadian policies governing such matters as cargo and passenger pre-clearance programs, law enforcement programs of all types, and immigration and refugee determination procedures.

The issues raised in this section, of course, apply largely to problems experienced in the exchange of goods. The cross-border exchange of services takes place to a significant extent without much government notice at the border. Some of it is exchanged electronically; some is eventually embedded in goods; and its extent is difficult to measure. Nevertheless, there are elements of border administration that are pertinent to trade in services, particularly professional services, entertainment services, and similar economic transactions that require the movement across the border of "natural" persons, i.e., real people. The CUSFTA and NAFTA include provisions to facilitate the temporary movement of business people, but these are a far cry from allowing the free movement of labour. Unlike the Schengen Agreement in Europe allowing for the free movement of people throughout the area covered by the arrangement, Canada and the United States still maintain stringent controls on the cross-border travel of each other's residents[58].

Given the extent of cross-border integration, the two governments have taken steps to address border congestion, but

---

criminals and other individuals who may pose a security risk. Although there is information sharing at the moment, it may need to be significantly upgraded and some of the information databases need to be combined and made available at the border. Information from law enforcement agencies, immigration agencies, the courts, and other institutions may need to be jointly accessible, at the border, in real time. There is need for much greater collaboration and better information management. Modern database management tools, as well as leading-edge networking software, can be deployed at the border and connected to main databases in Canada and the United States.

[58] See Michael Hart, *Is There Scope for Enhancing the Mobility of Labour Between Canada and the United States?* Paper prepared for Industry Canada, March 2004. Working Paper Series: Working Paper 2004 D-04, available at strategis.ic.gc.ca/epic/internet/ineas-aes.nsf/en/ra01934e.html.

with limited results to date. Efforts to make the border less intrusive and more efficient were integral to the 1996 *Shared Border Accord,* the 1999 *Canada-United States Partnership Forum,* the 2001 *Smart Border Accord,* and now the 2005 *Security and Prosperity Partnership*[59]. These initiatives, however, have been limited by the decision to work within the confines of existing legislative mandates. Furthermore, they assume continued need for current levels of border administration and thus are not aimed at eliminating or limiting the impact of the border, but at making border administration more efficient. Adding this effort to a broader commitment to negotiate a deep integration agreement would provide officials working on this file with the strategic vision they need to move beyond existing legislative mandates and provide them with greater scope to make useful trade-offs among competing priorities. Creating such a framework, investing in infrastructure and in technology (both at ports-of-entry and the corridors leading to such ports), and targeting resources toward pre-clearance programs for goods, vehicles, and people are critical components of any comprehensive effort at improving the management of the border and reducing its commercial impact. Ultimately, in our view, the objective should be to create a border that is considerably more open and less bureaucratic, within a North America that is more secure. If Canadians and Americans want a smarter and less intrusive border between them, they will also need to cooperate to create a more secure perimeter. The result should be a more open, more prosperous, and more secure continent.

---

[59] Pressure to be seen to be doing something disposes governments to artfully repackage earlier efforts in order to create new "announceables." Students of this phenomenon would do well, for example, to study the evolution of these four initiatives. Each promised concerted action at the level of the executive branch of government to address a series of border-related problems within existing legislative frameworks. All four shied away from any commitments that might lead to new treaty-level obligations that would require legislative approval. More may not have been politically feasible, but it is unrealistic to expect substantive results without a willingness to invest in more robust projects that might require legislative implementation.

Regulatory Cooperation

A key component to trimming border congestion lies in meeting the second challenge: reducing the impact of regulatory differences between Canada and the United States. As the Canadian Council of Chief Executives points out, "most of the administrative costs and delays at the border come not from the need to assess customs duties, but from myriad rules and regulations that are simply convenient for governments to handle at the border."[60] As Europeans learned, regulatory cooperation and reducing border formalities are two sides of the same coin. There may be a long tradition of pragmatic, informal problem solving between the regulatory authorities of the two federal governments, as well as among provincial and state governments, but the time has come to ask how much regulatory enforcement needs to be exercised at the border and how much can be exercised behind the border. More fundamentally, as regulatory cooperation and convergence proceed, the question becomes whether the two governments are ready to proceed to a more formal, treaty-based process of regulatory cooperation aimed at eliminating to the largest extent possible what has been characterized as the tyranny of small differences. By eliminating those differences, much of the rationale for border administration disappears. As well, it would increase the benefits of deepening integration.

The need to produce multiple versions of the same good, for example, can increase design and production costs, and prevent firms from enjoying the economies of scale that would flow from producing to satisfy a single globally accepted standard. An ever-growing range of goods have to be tested and certified to exacting standards and regulatory requirements before they can be sold. An equally exploding range of services faces onerous and often repetitive qualification and certification requirements. Compliance with different national rules, together with the repetition of redundant testing and certification of

---

[60] Canadian Council of Chief Executives, "New Frontiers: Building a 21st Century Canada-United States Partnership in North America." At www.ceocouncil.ca/en/view/?document_id=365&area_id=7.

140

products and providers for different markets, raises costs for manufacturers and providers operating in the North American economy. Additionally, complex and lengthy product- or provider-approval procedures can slow down innovation, frustrate new product launches, and operate to protect domestic producers and providers from foreign competitors. For small- or medium-sized firms, the cost of acquiring knowledge of and access to another country's regulatory regime can also effectively dissuade them from attempting to develop that market altogether.

While well-conceived regulations can be trade promoting and facilitating, regulatory divergence with the United States undermines Canadian competitiveness and results in lost investment. Recent research shows that the benefits of convergence between Canada and the United States are positive and significant. At the same time, Canada's regime, even allowing for important reforms that occurred over the last two decades, imposes significantly heavier burdens on the economy than that of the United States. If the burden of regulation in Canada had been the same as that of the United State, there would have been an average increase of investment in Canada of about US $1 billion annually. If the rate of change in the Canadian regime had been the same as that in the United States, the total investment would have been about $400 million higher resulting in an average of 30 percent more investment annually than the level that occurred. One consequence of such increased investment would have been a 6 percent increase in the research and development share of the GDP[61].

These impacts can be divided into two broad categories: those intended to discriminate in favour of local producers, and those that are the incidental result of regulations aimed at other objectives. The first represents the residual elements of traditional trade liberalization negotiations, and includes such measures as remaining tariffs, government procurement restrictions, trade remedy laws, and similar measures. The second involves a

---

[61] See Fidele Ndayisenya, "Economic Impacts of Regulatory Convergence Between Canada and the United States," In *Horizons*, 7:1 (June 2004), Canadian Policy Research Initiative.

wide range of measures that reflect the increasing complexity of modern economies and the response of governments to demands ranging from consumer protection to environmental stewardship and human rights. The trade and investment effects of the first could continue to be addressed with the traditional approach embedded in trade and investment liberalization agreements; the second would require higher levels of cooperation to identify those regulations that no longer serve any useful public purpose, those that can be implemented and administered on a basis that limits or eliminates the impact of differences, and those where differences are profound and important. Only the latter may need to continue to create any substantive barriers to trade and investment, but on a much more limited basis than is often the case today.

In the final analysis, however, many of these differences are marginal in their regulatory outcomes, particularly between Canada and the United States, but annoying and even dysfunctional in their economic impact. For companies exporting to multiple markets, the promise of "one standard, one test, accepted everywhere" has become increasingly more attractive[62]. Despite populist notions to the contrary, US regulatory requirements are often more stringent than those in Canada. To take one politically salient example, US responses to environmental degradation, from carbon emissions to water pollution, are often ahead of Canadian efforts. Notes George Hoberg: "as a result of policy integration through emulation, common science and technology, and shared values and politics, environmental policy in Canada and the United States has witnessed a substan-

---

[62] The OECD's Philip Wagner indicates how everyone would benefit from achieving this goal: "With harmonised standards and certification procedures, consumers can be confident that products sold throughout the global marketplace meet the same high safety standards everywhere. Manufacturers can avoid costly and unnecessary testing, and their innovative products will gain access to markets more speedily. Regulators can deploy increasingly scarce resources elsewhere, confident that products have been adequately tested and meet exacting requirements." Christopher Wagner, "Safe Products and Global Trade," *The OECD Observer,* No. 202 (October/November 1996), 16.

tial amount of convergence."[63] But, as Nancy Olewiler points out, Canada's "kinder, gentler route to improving the environment ... also means that Canada may not be moving as fast as it could toward reaching environmental targets."[64] A coherent program of cross-border cooperation is thus likely to strengthen Canadian regulatory outcomes, even one that will require Canada to do much of the heavy lifting and adjustment.

More to the point, bilateral regulatory convergence is more likely to involve adoption of best practices than reliance on the lowest common denominator. As an Industry Canada survey of Canadian regulators notes: "All of those surveyed indicated that their broad policy objectives were similar to those of their US counterparts. However, many stressed that differences in the respective systems of government and authorizing legislation complicate efforts to cooperate, effectively limiting what can be achieved without significant legislative changes." The same survey also indicated that "most cooperation takes place at the operational level."[65] At the same time, as the survey notes: without an external prod such as a trade negotiation, regulatory cooperation among those operationally responsible quickly grinds to a halt; without the involvement of regulators in the negotiations, however, the required objectives and means may not be well framed, leading to sub-optimal results.

The default option in addressing regulatory convergence between Canada and the United States has been to stay on the very Canadian path that has gradually emerged: cooperation if necessary but not necessarily cooperation. The results have not been uninteresting: Canadian jurisdictions align their regulatory goals and objectives with those of their US counterparts, work with US regu-

---

[63] George Hoberg, "Governing the Environment: Comparing Canada and the United States," in Keith G. Banting, George Hoberg, and Richard Simeon, eds., *Degrees of Freedom: Canada and the United States in a Changing World* (Montreal and Kingston: McGill-Queen's Press, 1997), p. 384.

[64] Nancy Olewiler, "North American Integration and the Environment," in Richard G. Harris, ed., *North American Linkages: Opportunities and Challenges for Canada* (Calgary: University of Calgary Press, 2003), p. 619.

[65] Industry Canada, *North American Regulatory Cooperation*, draft, February 2002, authors' files.

lators in many areas, but maintain sufficient regulatory autonomy to chart their own path. The result is two very similar but autonomous regulatory regimes involving extensive duplication and redundancy. The extent of regulatory convergence and cooperation is often influenced by bureaucratic agendas and preferences at the expense of broader goals from economic development to regulatory efficiency. This default position also avoids confronting the two related issues: the border and institutional capacity.

The External Advisory Committee on Smart Regulation EACSR) appointed by Prime Minister Chrétien concluded that this model was inadequate to address Canada's needs and recommended a proactive approach. It recognized that Canada is "enmeshed in a dense web of international relations," but wondered "whether the government's international regulatory activity is well aligned with national priorities and whether resources are being put to best use." It recommended that "the federal government should work to: achieve compatible standards and regulation in areas that would enhance the efficiency of the Canadian economy and provide high levels of protection for human health and the environment; eliminate small regulatory differences and reduce regulatory impediments to an integrated North American market; move toward single review and approval of products and services for all jurisdictions in North America; and put in place integrated regulatory processes to support key integrated North American industries (e.g., energy, agriculture, food) and provide more effective responses to threats to human and animal health and the environment."[66]

The Committee also recognized two policy traps that continue to appeal to some Canadians:

- Canadians can align their regulations with those in the United States to the extent they collectively judge it to be desirable on their own and do not need to complicate this process by tying it to a bilateral program; and

[66] External Advisory Committee on Smart Regulation, *Smart Regulation: A Regulatory Strategy for Canada (*Ottawa, Privy Council Office, 2004). At http: epe.lac-bac.gc.ca 100 206 301 pco-bcp committees smart_regulation-cf 2006-10-11 www.pco-bcp.gc.ca smartreg-regint en 08 rpt_fnl.pdf; pp. 17 & 22.

- Canadians should make a strategic judgment of where they want to be competitive, and then decide whether it is best to achieve that by being the same, being better, or being different.

While there is a superficial appeal to both points, experience suggests a unilateral approach is less likely to yield the desired result: reduce the impact of the border on Canadian trade and investment patterns. This goal will not be achieved in the absence of US confidence that Canada's regulatory regime is substantively equivalent to its own, a confidence that will require its active engagement. The existence of an agreed bilateral program, even one that may require Canada to adapt and adjust much more than the United States, has the additional clear advantage of bringing political pressure to bear on a process that would otherwise become too easily captive of bureaucratic agendas. The prime objective of such a program would not be to promote regulatory convergence for its own sake, but to enhance the performance of the Canadian economy by reducing barriers to reaping the full benefits of North American integration.

The government broadly accepted the recommendations of the EACSR and, in the context of the *Security and Prosperity Partnership* (SPP) adopted by the Presidents of the United States and Mexico and the Prime Minister of Canada in Waco, Texas in March 2005, took important steps to move the agenda along. It appointed a group in the Privy Council Office (PCO)[67] to pursue the path charted by the EACSR. Additionally, the federal government's Policy Research Initiative (PRI) was charged to consider ways and means to implement the EACSR recommendations. In his meeting with the two presidents in April 2006, Prime Minister Stephen Harper confirmed his government's commitment to the continued pursuit of the SPP agenda. What has emerged to date is a commitment to what might be characterized as accelerated incrementalism. The result is a higher level of awareness of regulatory developments in the

---

[67] As part of its effort to decentralize decision-making and reduce the PCO to its traditional role of policy coordination, the Harper government assigned work on regulatory reform to Industry Minister Maxime Bernier and his officials.

145

United States among Canadian policy makers leading to enhanced opportunities to align Canadian regulatory policy with developments in Washington. What is missing is a strong political commitment to regulatory cooperation and a plan to put it into effect. Not surprisingly, the pace in achieving regulatory rationalization has been glacial.

The current Canadian approach also appeals to American regulators, who have to date exhibited little appetite for more. The US decision-making system is extraordinarily resistant to centralized control and thus a very difficult target for more than piecemeal, regulator-to-regulator cooperation. The US President, for example, may appoint the Commissioners to the Securities and Exchange Commission, but once in office, they act fully independently of his direction.

Nevertheless, in both Washington and Ottawa, efforts at regulatory reform and streamlining have gained a growing number of adherents. Congress in 1980 established an Office of Information and Regulatory Affairs (OIRA) in the Office of Management and Budget (OMB) and successive presidents have, through Executive Orders, set out the basis for OIRA to provide systematic, centralized review and appraisal of all federal regulations. Much of this effort has been coordinated with broader international initiatives, particularly at the OECD. Canadian efforts parallel those in the United States. Since 1978, Canadian federal regulatory activity has been subject to a constant, comprehensive, centralized process of review, housed initially in the Treasury Board and subsequently in the PCO, with a view to eliminating duplication and redundancy and promoting best international practice. The guiding policies developed in both capitals for rule making and review are remarkably similar in tone and intent and reflect the high level of ongoing discussions at the OECD and bilaterally[68]. A sound foundation has, therefore, been created for a more formal program of cross-

---

[68] See, for example, "Government of Canada Regulatory Policy" at http: www.tbs-sct.gc.ca ri-qr ra-ar docs publications regulatory_policy_e.pdf and United States. "Economic Analysis of Federal Regulations Under Executive Order 12866." At http: www.whitehouse.gov omb inforeg riaguide.html

border regulatory cooperation and even coordination. To go to the next level, however, the two governments would need to adopt a program leading to an enforceable agreement, and the institutional capacity to make it work[69].

Governments must think carefully about any initiatives that may compromise their ability to discharge their responsibility for the security and well-being of their citizens. Canadian experience in negotiating international rules and pursuing regulatory cooperation, both multilaterally and bilaterally, suggests that there is no inherent conflict between these responsibilities and such rulemaking and cooperation. Nevertheless, vested interests can mount emotional campaigns questioning the extent to which regulations made jointly with others can respond to Canadian responsibilities. Fortunately, it is not difficult to refute such claims. Canadians, for example, routinely travel in the United States, comfortable in the reliability of US safety regulations. They eat and drink in the United States on the same basis as they do at home. If they are sick, they often can and do rely, at considerable expense, on US medical advice and US-approved drugs. From almost any perspective, Canadians have few if any qualms about the goals and efficacy of US regulations when in the United States. There are few other countries about which Canadians routinely exhibit such confidence. The reason is simple: Canadian and US regulatory regimes are, in almost all respects, closely aligned. The differences are matters of detail that may matter to individual regulators, but have little impact on residents in either country.

Initially, the two governments could build confidence and gain experience at the federal level, but given the federal structure of the two countries, the sooner they engage provincial and state regulatory authorities in a similar process of mandatory information exchange and consultations, the sooner the two countries would arrive at a "North American" approach to meeting

---

[69] For more detail on what this would involve, see Michael Hart, "Steer or Drift? Taking Charge of Canada-US Regulatory Convergence," *C.D. Howe Institute Commentary*, No. 229 (Toronto, March 2006). This *Commentary* also includes a discussion of the experience in Europe in transforming its customs union into an integrated market.

their regulatory goals and objectives. Because of the large number of jurisdictions involved, this is an area that would require some creative decision rules as well as institutions to make them work. Fortunately, as at the federal level, extensive regional networks of collaboration already exist between Canadian and US regulators. Any successful federal strategy on economic integration and regulatory convergence would need to both complement and take advantage of these existing cross-border institutions.

Institutional Capacity

Integral to any progress in addressing the governance of deepening integration is the need to build sufficient institutional capacity and procedural frameworks to reduce conflict, and provide a more flexible basis for dynamic rule-making and adaptation for the North American market as a whole. It may well be necessary to consign traditional aversion to bilateral institution building to the dustbin and look creatively to the future. While the European model of a complex supranational infrastructure may not suit North American circumstances, there are lessons Canadians and Americans can learn from the EU experience.

The deep integration described above has occurred in the historical absence of an institutional infrastructure for managing this complex, multifaceted relationship. Unlike other bilateral relationships enjoyed by both Canada and the United States, there is no body to provide political or policy oversight of the relationship, no regular meetings between heads of government or foreign or trade ministers, no formal structure of committees looking at the relationship in a coherent and coordinated manner[70]. The absence of formal structure results from a determined

---

[70] As former Canadian ambassador to the United States, Allan Gotlieb, observes, "the world's largest bilateral economic relationship [is] managed without the assistance of bilateral institutions and procedures." "A Grand Bargain with the US," *National Post,* 5 March 2003. This paucity of institutions stands in stark contrast with a veritable cornucopia of institutional relationships with the European Union including biannual meetings of the Prime Minister, the President of the European Commission and President (in office) of the Council as well as a host of ministerial committees, official working groups etc. See www.dfait-maeci.gc.ca/canada-europa/mundi/menu-en.asp.

148

and largely successful effort to treat issues in the relationship vertically, rather than horizontally, and to build firewalls to prevent cross linkages. In part, this method of management derives from Canadian fears that, as the smaller partner, Canadian interests would be overwhelmed in any more formal relationship. In part, it originates in the US system of governance that makes coherence and coordination in both foreign and domestic policies extraordinarily difficult to achieve on a sustained basis.

The institutional gap is filled by inspired *ad hocery.* The inter-connected natures of the Canadian and American economies virtually require Canadian and US officials to work closely together to manage and implement a vast array of similar but not identical regulatory regimes from food safety to refugee determinations. Officials and, in some cases, ministers have developed a dense network of informal cooperative arrangements to share information, experience, data, and expertise with a view to improving regulatory outcomes, reducing costs, solving cross-border problems, implementing mutual recognition arrangements, establishing joint testing protocols, and more. On any given day, dozens of US and Canadian officials at federal, provincial, and state levels are working together, visiting, meeting, sharing e-mails, taking phone calls, and more. Virtually all of this activity takes place below the political radar screen. Little of it is coordinated or subject to a coherent overall view of priorities or strategic goals. Some of it is mandated by formal agreements ranging from the NAFTA to less formal memorandums of understanding. More importantly, much of this activity is the natural result of officials with similar responsibilities and shared outlooks seeking support and relationships to pursue them. This activity also reinforces, subtly and indirectly, the deepening integration of the two economies. The NAFTA and similar arrangements mark efforts by governments to catch up with these forces of silent integration and provide appropriate and facilitating governance.

The focus should be upon the functions that need to be performed for the efficient governance of deepening integration, rather than the creation of new institutions where none is needed. To some extent, these could be met by making creative

use of existing Canada-US cooperative arrangements, by investing officials in agencies on both sides of the border with new responsibilities, or by building on existing models that have worked well.

The two governments could, for example, stipulate that the Canadian Border Services Agency and the US Customs Service coordinate their efforts to ensure efficient administration of third-country imports. Similarly, an appropriate understanding could be reached requiring the Canadian Department of Transport and the US Department of Transportation to coordinate their efforts to ensure highway safety; before enacting any new rules and regulations, for example, mandatory coordination efforts would focus on ensuring compatible outcomes and mutual recognition of each other's approaches to the same problem. A good basis for this kind of cooperation already exists in both informal networks among officials, and in the relatively minor differences in regulatory approach. What is missing is an agreed mandate to resolve differences and a more formal institutional framework with authority to ensure mutually beneficial outcomes. Establishing a bilateral commission to supervise efforts to establish a more coordinated and convergent set of regulations governing all customs or transportation matters could prove critical to providing the necessary momentum and political will.

In both countries, labour mobility is hampered by provincial and state labour laws and delegated professional accreditation procedures. The NAFTA put in place a modest process to permit temporary entry for business and professional visitors and mutual recognition of professional accreditation. The latter has been hampered by the conflict of interest inherent in a system of self-regulation. As the EU learned, a more centralized approach was required to overcome conflicts of interest and bureaucratic inertia. From architects and accountants to doctors and dentists, there remains considerable scope for enhancing mutual recognition arrangements. An important step toward breaking the logjam would be to appoint a bilateral task force to develop model mutual recognition arrangements for consideration by state and provincial accreditation bodies.

Much can be achieved on the basis of existing networks of cooperation, with the addition as necessary of specific joint or bilateral commissions in instances where existing networks are inadequate. More will be achieved, however, if the two governments commit to the establishment of a limited number of bilateral institutions with a mandate to provide them with the necessary advice and information to effect a more integrated North American approach to regulation. An independent Canada-US Secretariat with a mandate to drive the agenda and report annually to the President and Prime Minister on progress could, for example, prove critical to overcoming bureaucratic inertia. Similarly, a Joint Advisory Board to the President and Prime Minister could contribute some creative drive to the development of new bilateral initiatives. As numerous studies have demonstrated, regulatory agendas are prone to capture, geared to serving the narrow interests of regulator and regulatee. Bilateral initiatives limited to regulatory authorities are unlikely to prove immune from this reality. Regular review by an independent advisory board of progress in implementing a bilateral program of "guided" regulatory convergence could thus prove a valuable addition in keeping the program focused on broader objectives.

The Mexico Question

In concluding the NAFTA in 1993, Canada and the United States opened a new era in their relations with Mexico. The NAFTA stands as testimony to the belief that the North American community involves three nations and that the destiny of all three nations is inextricably intertwined. Broad consensus is in our view emerging, however, that for the moment the Canadian challenge is to elaborate a bilateral rather than a trilateral agenda [1]. Successful implementation of the NAFTA ushered in

[1] For some analysts, particularly those with legal training, the NAFTA now constitutes the indispensable foundation from which all future bilateral or trilateral trade and investment negotiations must proceed. Perhaps, but it is also possible to view the NAFTA as an agreement whose time came but is now largely gone. Between Canada and the United States, it is

expectations of closer trilateral relations, but the reality is that the NAFTA provides a common framework of rules to govern two robust and rapidly evolving relationships: between the United States and Canada and between the United States and Mexico. Canada-Mexico relations remain at a much less advanced stage, while the priority issues between Canada and the United States and Mexico and the United States are not of the same order. There are issue areas where there may already be scope for advancing toward trilateral rules and institutions, e.g., surface transportation, while in other issue areas much more will be gained from parallel bilateral efforts, e.g., energy. To that end, the three governments may wish to consider pursuing ways and means to network any bilateral and or efforts and ensure that any success at the bilateral level ultimately feeds into trilateral goals and aspirations[72].

---

now fully implemented. Most of its provisions would be very difficult to undo. Its main continuing importance lies in the dispute settlement provisions set out in chapters 11 and 19; even chapter 20 has been largely superseded by the WTO Dispute Settlement Understanding. We do not mean to suggest that we can now abrogate the NAFTA, but rather that like the 1935 and 1938 Reciprocal Trade Agreements, the 1947 GATT, and the 1965 Autopact, the NAFTA's objectives have been largely realized. Any successor agreement may need to incorporate some of its elements (e.g., chapters 11 and 19), but need not be trilateral. Should Canada and the United States proceed to an agreement that moves well beyond the contours of conventional trade agreements, the precedential value of the NAFTA will be minimal.

[72] In "A Trilateral Mirage: A Tale of Two Americas," a paper prepared for the Canadian Defence and Foreign Affairs Institute in June 2003, Jean Daudelin comes to a less accommodating conclusion: "Canada's bilateral relationship with the United States is vital and its management should not be cluttered by the massive complexity of Mexico-US affairs. ... Canada's relations with Mexico ... will remain marginal to the country's core interests." Jeffrey Schott, "Prospects for North American Economic Integration: An American Perspective Post-9/11," *Art of the State II: Thinking North America: Prospects and Pathways* (Montreal: Institute for Research on Public Policy, 2004), on the other hand, concludes that US political sensitivities make anything other than a trilateral initiative unrealistic. Discussions with US officials suggest that US political sensitivities are more nuanced, providing scope for differentiated approaches. The difficulties of a trilateral approach that considers both economic and security issues are also well illustrated by the paper prepared for The Art of the State II Conference by Stéphane Roussel and Athanasios Hristoulas,

The Security and Prosperity Partnership of 2005 and renewed in 2006, the official template for the discussion of emerging issues among the three federal governments of North America, specifically recognizes the reality of a "two-speed" approach. Future work on bilateral issues between Canada and the United States, therefore need not become hostage to whether or not Mexico is interested in, or has the capacity to pursue, a specific matter.

The Sovereignty Dimension

Canadian trade policy debates are virtually unique in their elevation of sovereignty as a critical issue of public policy[73]. In the great debates of the past over free trade, notably the elections of 1891, 1911, and 1988, supporters of free trade with the United States found themselves under sustained attack on the grounds that the sacrifice of Canadian sovereignty was too heavy a price to pay. Any new trade negotiations with the United States is certain to revive this old, if increasingly quaint, discussion along a number of fronts.

One argument, advanced by the Senate Foreign Affairs Committee, is that concluding any deeper integration agreement, such as a customs union, would rob Canada of the capacity for setting tariffs on trade with third countries[74]. The reality is that Canada has given up tariff autonomy through the progressive binding of its tariff in the GATT and WTO negotiations as well as the declining utility of the tariff as a policy instrument. A variant on this theme is the claim that a new agreement would prevent the pursuit of distinct industrial, energy,

---

"The Quest for trilateral security in North America" (Montreal: Institute for Research in Public Policy, 2003).

With the exception of defence policy issues, where public discussion often veers into dead-end tangents divorced from the iron realities of protecting the two countries from external threats.

[74] Senate Standing Committee on Foreign Affairs, *Uncertain Access: The Consequences of US Security and Trade Actions for Canadian Trade Policy* (June 2003) at parl.gc.ca.

immigration, and environmental policies[75]. This point also lacks substance. The scope for industrial and energy initiatives is already severely constrained by WTO rules on subsidies in the former and by the NAFTA rules in the latter. As regards issues such as immigration and environment, it is far from clear how creating an agreement to enlarge the benefits of economic integration would limit Canadian policy.

A second line of criticism is that a customs union agreement would tightly tether Canadian foreign policy to that of the United States. During the free trade negotiations in the 1980s, critics charged that the United States would threaten Canadian access to its market provided for under the agreement unless Canada embraced US foreign policy goals. Such criticism should be dismissed on a number of grounds. The United States has no modern history of withdrawing from trade agreements or reducing access to its market provided thereunder for foreign policy reasons. Second, the vulnerability of Canada to such a remote probability arises not from the existence of bilateral agreements, but from the deep integration of the two economies, an integration occurring principally because of the natural dynamics of economic integration, driven by the daily economic choices of individuals and firms and fostered by past multilateral and bilateral trade agreements. The only way to mitigate this vulnerability to the United States is to reverse the course of economic integration and accept the incalculable economic costs of such action as a price of insulation from US influence[76].

---

[75] For a sample of these views, see Andrew Jackson, "Why the 'Big Idea' is a Bad Idea: A Critical Perspective on Deeper Economic Relations with the United States" (Ottawa: Canadian Centre for Policy Alternatives, 2003). See also Stephen Clarkson, "Time to break free (trade)," *Globe and Mail,* 27 September 2002 and Peter Newman, "Beware of freer trade," *Maclean's,* 2 December 2002. A more thoughtful version of this view can be found in a series of columns by David Crane in the Toronto Star, August 9, 13, and 16, 2003. Support among Canadians for such a defensive attitude toward the United States, however, has steadily declined.

[76] Canadians have demonstrated that they have become increasingly comfortable with their proximity to the United States and with the pragmatic pursuit of better ways to manage deepening integration. See the discussion of

154

A third criticism is that a customs union would further erode Canada's historical multilateral vocation and undermine Canada's ability to participate in multilateral negotiations.[77] This argument reflects the continuing allure of multilateralism as the central organizing principle of Canadian trade and foreign policy. In fact, an objection to new bilateral arrangements on the grounds of protecting Canada's multilateral heritage is not only historical revisionism, it confuses ends with means[78]. Multilateral rule making and institution building have proven effective means for Canada to pursue its trade objectives but have never impeded the pursuit of trade or foreign policy objectives by other means. Canada has been prepared to look to bilateral rules and institutions when these are available and better suited to achieve the policy objective sought. Both bilateral and multilateral strategies need to be judged on their ability to satisfy Canadian needs and interests. To forego benefits available in a bilateral arrangement in order to uphold the multilateral ideal would make a nullity of coherent policy making.

A fourth criticism is that a deep integration agreement would ignite a "race to the bottom," i.e., to a relentless effort by governments to attract foreign investors and retain domestic investors by reducing regulatory norms and expectations. There is little evidence to support this charge. Indeed, there is a preponderance of evidence pointing in exactly the opposite direction. As societies become more prosperous—one of the most important impacts of globalization and of deepening integration—the demand for regulations to enhance the quality of life increases. The explosion of government regulatory activity to address environmental, human rights, safety, and other issues provides com-

---

recent polling in Hart, "A New Accommodation with the United State: The Trade and Economic Dimension."

[77] See, for example, Robert Wolfe, "See You in Washington? A Pluralist Perspective on North American Institutions," IRPP *Choices*, 9:4, at irpp.org.

[78] Brian Tomlin and Bruce Doern, *Faith and Fear: The Free Trade Story* (Toronto: Stoddart, 1991), debunk this cherished myth of Canadian foreign policy elites.

pelling evidence of the gap between rhetoric and reality. In the other direction, the impact of regulatory convergence and regulatory cooperation has been repeatedly to raise the bar by establishing international benchmarks of minimal performance and best international practice. As argued above, bilateral regulatory convergence is more likely to involve adoption of best practices than a lowering of standards to the lowest common denominator.

Conclusions

The Canadian and US economies have become intertwined in response to demands by Canadians and Americans alike for each other's products, services, capital, and ideas, creating jobs and wealth across many sectors and accelerating the forces of mutually beneficial integration. The strategic challenge facing Canadians today is whether they want their government to help or hinder accelerating cross-border economic integration and social and political interaction. Whatever the homilies about the value of independence, there is no sentiment that the government should interfere in private business and investment decisions to change the logic of resources, geography, and private choice that underpin economic integration.

The trade policy of the post-war years, grounded in the theories of well-established international trade theory and pursued on the basis of the politically pragmatic strategy of mercantilist bargaining, proved critical to underwriting the first stages of cross-border integration. The framework of rules and institutions developed over the past seventy years have worked well to facilitate and govern a process of "silent," market-led integration. Both the theory and the practice responded well to the issues that concerned the dominant economic players: large, vertically integrated, multinational companies with operations on both sides of the border focused on the production and exchange of physical goods. While such companies remain important players in North America, the development of much more fragmented production strategies, the ability to disperse production much more widely around the world, the emergence of new security threats, and the reality of a much wider range of cross-

border transactions, all point to the need to look at a new set of policy issues that threaten to disrupt the mutually beneficial process of cross-border integration.

The continued presence of a heavily administered border, of similar but differentiated regulatory regimes, and inadequate institutional capacity to solve problems, now undermines the ability of firms and individuals alike to reap the full benefits of deepening cross-border integration. New developments in trade and industrial organization theory help us to understand the nature of new production and exchange patterns and underline the extent to which subtle differences in regulatory regimes and undesirable border complications can frustrate corporate efforts to organize production and exchange as efficiently as possible. In a world where firms have many more choices about what to produce and where, the smaller partner in a deeply integrationist relationship is particularly vulnerable to the impact of border delays and regulatory differences. In these circumstances, the Canadian government is well-advised to invest in efforts that will bring the framework of rules governing cross-border exchange into line with commercial and economic reality.

The Consequent Policy and Research Agenda

Pursuing the policy issues outlined above would be facilitated if analysts, both inside and outside government, developed further some of the ideas set out in this paper and the data and analysis upon which it is based. To that end, we believe further work along the following lines would make a material contribution to informed policy discussion:

- More detailed assessments of the extent of fragmentation and integration involving firms and facilities with cross-border interests in Canada and the United States.
- The extent of investment and other corporate ties linking production in the two countries.
- An inventory of the issues critical to business and investment in making cross-border trade and investment function to its full potential.

- An inventory of the customs and related functions performed at the border that can be performed behind the border, facilitating pre-clearance and similar strategies.
- The relationship of formal border clearance procedures to the interdiction of criminal, terrorist, and other unlawful cross-border activities.
- An inventory of Canadian and US regulatory regimes with serious implications for cross-border trade and investment, and the extent of differences in each of these regimes.
- An inventory of the extent and nature of cross-border networks of decision-makers and an assessment of their effectiveness in addressing cross-border problems.
- An assessment of the European experience in governing deepening integration and the extent of its applicability to North American experience.

We list these as indicative of the range of interesting issues that warrant timely investigation and consideration in building the necessary intellectual capital to address the issues raised in this paper. Other may have additional ideas. What's important is that the analytical and policy communities devote the necessary intellectual resources to thinking through what would be involved in making the North American economy function even better in meeting the needs of Canadians and Americans alike.

Part III

Regional Trade Issues

US–Canadian Trade and US State-Level Production and Employment: An Update

Laura M. Baughman and Joseph Francois[*]

Abstract

Using a computable general equilibrium (CGE) model belonging to the class of multi-region CGE models commonly used to estimate the economy-wide and the sector-specific impacts of trade policy changes, we estimate the impact on US and state output of changes in the US–Canada trading relationship, and their resulting impacts on US jobs. We find that trade with Canada provides tangible and important economy-wide employment and income benefits to the United States and to every US state. Total trade with Canada—of goods and services, and exports as well as imports—generated US output worth $327 billion in 2005, or 2.6 percent of GDP. Output supported by total trade with Canada also support US employment levels. We estimate that trade with Canada supported more than 7.1 million net US jobs, or 4.1 percent of total US employment in 2005. Every US state registered net positive job gains from trade with Canada.

Introduction

Both the United States and Canada have experienced their shares of debates about the costs and benefits of trade. From the Canada–US Free Trade Agreement (CUSFTA) to the North American Free Trade Agreement (NAFTA) to bilateral dis-

---

[*]Laura Baughman is president and managing director of Trade Partnership Worldwide, LLC. Joseph Francois is professor of economics with the Johannes Kepler Universität Linz, a fellow of the Tinbergen Institute, a research fellow with the Centre for Economic Policy Research, and a member of Trade Partnership Worldwide. The views expressed in this chapter are those of the authors.

159

putes, the identification of the costs of trade, and particularly the costs of the bilateral trading relationship, seems to over-shadow discussion of the benefits of the trading relationship.

Notwithstanding the history of bilateral trade disputes, the US–Canada trading relationship is clearly a case where both sides benefit in important ways. Bilateral trade between the two neighbours has been growing steadily for decades. The increasingly integrated nature of the two economies—thanks to CUS-FTA and NAFTA—now has developed into a generally comfortable integrated relationship. Nevertheless, arguments persist that this integration has cost hundreds of thousands of workers their jobs. Indeed, the jobs debate overshadows the otherwise positive assessment of the impact of increased trade.

With the importance of the trade–jobs question as motivation, in this paper we update and expand on earlier research that quantifies the US employment impact of US–Canada trade. (Francois and Baughman 2004). Our earlier research found that in 2001, cross-border trade (exports and imports) in goods supported 5.2 million US jobs. The current research updates that employment estimate to 2005, and assesses the net job impact in the United States of cross-border services trade[1]. As in our earlier research, we break down the job estimates by state. We begin with an overview of US–Canada trade in goods and services, and then present our estimates of the number of US jobs supported by trade with Canada in 2005. We conclude with a summary of what our results mean for US trade policy. Appendix A details our methodology for estimating the US job impacts of US–Canada goods and services trade.

The US–Canada Trading Relationship

Canada is, not surprisingly, one of the United States' leading trading partners. It is the largest foreign market for US goods, and the largest source—exceeding even China—of US imports. It is the third most important market for US services exports, and the fourth largest source of US services imports (see Table 1,

---

[1]Because of the addition of services trade to the modelling exercise, results from the 2001 analysis are not comparable to the results of this analysis.

which reports the top ten markets for US exports and imports of goods, and US exports and imports of services).

While clearly much of this tendency to trade with each other is owed to geography, also important has been the success of efforts both countries have undertaken to reduce and eventually eliminate barriers to trade between them. CUSFTA went into effect on 1 January 1989, aiming to eliminate bilateral tariffs and many non-tariff barriers in most sectors of merchandise trade within 10 years. NAFTA replaced CUSFTA on 1 January 1994. By that time, most US–Canada trade was already duty-free thanks to CUSFTA. By 2005, virtually all bilateral trade in goods and services was trade-barrier free.[2]

Table 1: Leading US Trading Partners, 2005
(billions of US dollars)

	Goods		Services	
	Exports	**Imports**	**Exports**	**Imports**
Total, World	$894.60	$1,677.40	$360.50	$280.60
Canada	*212.2 (1ˢᵗ)*	*293.3 (1ˢᵗ)*	*32.8 (3ʳᵈ)*	*22.5 (4ᵗʰ)*
United Kingdom	*37.6*	*50.5*	*45.7 (1ˢᵗ)*	*36.8 (1ˢᵗ)*
Japan	*53.3 (3ʳᵈ)*	*138*	*42.5 (2ⁿᵈ)*	*23.8 (2ⁿᵈ)*
Mexico	*120.3 (2ⁿᵈ)*	*172.1 (3ʳᵈ)*	20.6	14.9
Germany	*33.6*	*84.6*	20.3	*25.6 (3ʳᵈ)*
France	*22.3*	*33.8*	13.2	12.9
Korea	*27.1*	*43.8*	*11*	7.9
Switzerland	10.7	13	9.5	*11.4*
China	*41.8*	*243.5 (2ⁿᵈ)*	9.1	6.5
Netherlands	*26.3*	14.8	9.1	8.1
Bermuda	0.5	nil	n.a.	14.1
Taiwan	*22.5*	*34.8*	7.8	6.7
Venezuela	6.4	*34*	2.6	0.6

Italics — country is among the top ten for that type of export or import.

n.a. = not available.

Source: US Department of Commerce, Bureau of Economic Analysis

---

*This is not to suggest that there are no restrictions on bilateral trade. Certain sensitive sectors in both economies continue to face bilateral trade restrictions, while antidumping and countervailing duty actions continue to result in disruption of trade for individual products.

While US goods trade with Canada has been growing over the years, Canada's importance as a trading partner in the post-FTA period has been comparatively stable. In the period 1995–2005, Canada's share of total US exports ranged from 22 to 24 percent, with a modest upward trend over the period. Canada's share of total US imports, on the other hand, declined more or less steadily over the period as lower-cost foreign suppliers, particularly in China, claimed larger shares of the US import market. Canada enjoyed a growing trade surplus with the United States over this 10-year period; however, this increase generally mirrored the total US trade balance, as Canada's share of the total US trade balance has generally averaged about 10 percent.

Table 2: US Goods Trade with Canada, 1995–2005

	Exports	Imports	Balance
	Value (billions of US dollars)		
1995	$127.4	$146.9	-$19.5
1996	134.3	158.5	-24.3
1997	151.9	170.1	-18.2
1998	156.7	175.8	-19.1
1999	166.7	201.3	-34.6
2000	178.9	233.7	-54.8
2001	163.3	218.7	-55.5
2002	160.9	211.8	-50.9
2003	169.8	224.2	-54.4
2004	190.0	259.0	-69.0
2005	212.2	293.3	-81.1
	Share of Total US (percent)		
1995	22.1	19.6	11.2
1996	21.9	19.7	12.7
1997	22.3	19.4	9.2
1998	23.3	19.2	7.7
1999	24.4	19.5	10.0
2000	23.2	19.1	12.1
2001	22.4	19.1	13.0
2002	23.6	18.2	10.6
2003	23.8	17.8	9.9
2004	23.5	17.6	10.4
2005	23.7	17.5	10.4

Source: US Department of Commerce, Bureau of Economic Analysis

While Canada enjoys a goods trade surplus with the United States, the United States enjoys a services trade surplus with Canada (Table 3). Trends in US services exports and imports with Canada mirror those of goods, but on a much smaller scale. Canada's share of total US services exports has been increasing since 1998, while its share of total US services imports has been declining over that period.

Table 3: US Services Trade with Canada, 1995–2005

	Exports	Imports	Balance
	Value (billions of US dollars)		
1995	$18.1	$11.2	$6.9
1996	19.6	12.6	7.0
1997	20.6	14.0	6.6
1998	19.6	15.3	4.3
1999	22.8	16.4	6.4
2000	24.7	18.0	6.7
2001	24.5	17.5	7.0
2002	25.1	18.0	7.1
2003	27.4	19.5	7.9
2004	29.7	21.1	8.6
2005	32.8	22.5	10.3
	Share of Total US (percent)		
1995	8.2	7.9	8.8
1996	8.2	8.2	8.2
1997	8.0	8.4	8.2
1998	7.5	8.5	7.9
1999	8.1	8.2	8.1
2000	8.3	8.1	8.2
2001	8.6	7.9	8.3
2002	8.6	7.8	8.2
2003	9.1	7.8	8.5
2004	8.6	7.3	8.0
2005	8.6	7.2	8.0

Source: US Department of Commerce, Bureau of Economic Analysis

Overall, US trade with Canada is huge and growing. Total trade (goods plus services, exports plus imports) reached $561 billion in 2005 and reflected average annual increases of 6.5 percent over the previous decade (Table 4). Similarly, total

goods trade (exports plus imports) has been growing at an average annual rate of 6.5 percent, and total services trade by even more: 6.7 percent per year.

Table 4: Total* Goods and Services Trade, 1995–2005
(billions of US dollars and percent)

	Total Goods	Total Services	Total Trade	Goods' % Share of Total
1995	$274.3	$29.3	$303.6	90.4
1996	292.8	32.2	325.0	90.1
1997	322.0	34.6	356.5	90.3
1998	332.6	34.9	367.4	90.5
1999	368.0	39.2	407.2	90.4
2000	455.3	42.8	455.3	90.6
2001	424.0	42.0	424.0	90.1
2002	415.8	43.1	415.8	89.6
2003	441.1	46.9	441.1	89.4
2004	499.8	50.8	499.8	89.8
2005	505.5	55.3	560.8	90.1

*Exports plus imports.
Source: Bureau of the Census

From a sectoral perspective, many categories of goods show up as both leading exports to and imports from Canada, suggesting co-production between producers in both countries (Table 5). This co-production is most obvious in the case of the motor vehicle sector, where the two countries' auto sectors have been integrated for many years. At the same time, Canada is an important source of raw materials to the US market, most notably petroleum (mineral fuels, which took the lead in 2005). Also important are wood and wood products.

In services trade, transportation-related services exports and imports are roughly comparable in size, although more Canadian passengers travel to the United States than US passengers to Canada. Trade in other private services trade has been the main source of growth in services trade, with US exports to Canada doubling over the period and US imports from Canada rising by a factor of more than 2½. The United States has a significant and growing bilateral surplus in royalties and license fees.

164

Table 5: Leading Sectors in US Trade with Canada, 1995, 2000, 2005 (billions of US dollars)

Goods Exports	1995	2000	2005
Vehicles (HS 87)	$25.7	$32.8	$40.9
Non-electrical machinery (HS 84)	21.9	30.6	30.9
Electrical machinery (HS 85)	12.8	18.0	13.9
Plastics (HS 39)	4.4	6.9	9.4
Iron and steel (HS 72 & 73)	3.9	5.8	8.7
Mineral fuels (27)	1.4	2.6	8.1
Precision instruments (HS 90)	3.7	5.8	5.3
Paper, paperboard, paper pulp (HS 48)	2.5	3.7	4.3
Organic chemicals (HS 29)	1.6	2.2	3.3
Rubber and rubber products (HS 40)	1.9	2.8	3.0
Goods Imports			
Mineral fuels (HS 27)	$13.6	$31.4	$65.4
Vehicles (HS 87)	40.7	56.1	61.7
Non-electrical machinery (HS 84)	13.4	18.8	19.7
Wood and wood products (HS 44)	7.1	10.8	14.2
Electrical machinery (85)	6.9	16.9	10.8
Plastics and products (HS 39)	3.8	6.7	10.5
Paper, paperboard, paper pulp (HS 48)	9.0	10.1	10.4
Aluminium and aluminium products (76)	3.9	4.5	6.9
Aircraft (HS 88)	1.5	4.7	6.0
Furniture (HS 94)	1.2	5.3	5.8
Services Exports			
Transportation-related services	$9.8	$11.4	$14.6
Travel	*6.2*	*7.2*	*9.0*
Passenger fares	*1.3*	*1.6*	*2.6*
Other transportation	*2.3*	*2.6*	*3.0*
Other private services	6.7	10.7	13.4
Royalties and license fees	1.4	2.8	4.4
Services Imports			
Transportation-related services	$7.1	$10.7	$11.6
Travel	*4.3*	*6.2*	*7.0*
Passenger fares	*0.3*	*0.8*	*0.3*
Other transportation	*2.5*	*3.7*	*4.3*
Other private services	3.7	6.0	9.5
Royalties and license fees	0.2	1.0	0.8

Source: Bureau of the Census

The US–Canada Economic Relationship: What It Means for US Output and Jobs

The economic impacts of trade are one of the biggest concerns of policy-makers on both sides of the border. Polling results suggest that large percentages of the American public believe that trade expansion, and particularly increases in trade deficits, result in domestic job losses. This belief is so longstanding and prevalent that it is widely accepted as fact, and often left unchallenged in political debates.

The data actually support the opposite conclusion: trade (both exports and imports) creates output, which is job-supporting. Because of the role of Canadian inputs in integrated production processes in the United States, increased trade—including increasing imports—contributes to increased domestic output and related jobs. This includes manufacturing jobs.

Our earlier analysis of the US output and job impacts of US–Canada goods trade found that cross-border trade in goods in 2001 supported approximately $162 billion in US economic activity and 5.2 million net jobs (job gains net of job losses). These estimates understated the true value of the US–Canada trading relationship to the United States for two reasons. First, it measured the impact during a recession year, when trade flows were lower than they otherwise would have been had both economies been healthier. Goods trade dropped in 2001 from 2000 levels, and did not recover for several years, until 2004 (see Table 4). Second, the earlier research looked only at the impacts of cross-border trade in goods, rather than goods and services. Given the importance of services to both economies, the output and employment impacts of total trade would necessarily be understated.

In this paper we examine the output and employment impacts of the bilateral trading relationship for a year (2005) that more accurately reflects the robust nature of that relationship, and include total services trade as well. We use the same methodology we followed in our earlier research: a computable general equilibrium (CGE) model that examines the up- and downstream impacts of trade on the US economy.

A Brief Description of the Model

CGE models are commonly used today to estimate the economy-wide and the sector-specific impacts of a trade policy change. A global model in wide use today is the Global Trade Analysis Project (GTAP) model. Working with a version of this model (with the modifications and updates described below), we estimate the impact of US–Canada trade on US production, consumption, trade, prices and welfare.

Briefly, we have updated our core dataset (version 6.2) from 2001 to 2005, and focused on 14 sectors, four regions (the United States, Canada, Mexico, and the rest of the world), and fixed capital: in other words, our analysis is static. The model's structure assumes perfect competition and constant returns to scale. More details are provided in Appendix A. These structural features are appropriate given the current application.

To estimate the impacts on the United States in 2005 of exports and imports of goods and services, we posit the following counterfactual: suppose those exports and imports were simply eliminated as the result of the imposition of a prohibitive tariff on US imports from Canada, and the simultaneous imposition of a prohibitive tariff on Canadian imports from the United States. The losses in US output provide a measure of the opposite: the gains in US output linked to trade.[3] We then take these

---

[3]It is important to note that these estimates show what the level of US output and employment would have been in 2005 if US–Canadian bilateral trade were reduced to zero, with the rest of the world continuing and filling in where possible for the lost US–Canada trade. For example, in the counterfactual scenario some US imports of lumber for the housing sector from Canada would have been replaced by imports of lumber from other countries, and the impact would be felt in terms of the higher cost of the alternative sources of inputs. Accordingly, the reported estimates show the gains that the United States makes in jobs and output from being able to trade directly with Canada as well as with the rest of the world. There would, of course, be greater impacts of a full border closure (i.e., one that reduced to zero imports from all trading partners and exports to all trading partners). In addition, a permanent trade-related border closure (simply stopping all trade) can be quite different in effect from one limited temporarily to physical movement of persons and goods (and not, for example, electronic services).

national estimates (detailed by sector) and distribute them to states according to published sector-specific data for state output. Finally, we compute the jobs related to that output using state- and sector-specific production-to-employment ratios.

Our approach examines the impact of bilateral trade on US output and employment by accounting for the effects of both total exports and of total imports, rather than simply looking at the impact of net flows (the so-called trade deficit[4]). This approach better permits us to capture the full contribution of trade to the efficiency of US output and to employment. Exports and imports both support jobs directly, jobs tied to manufacturing goods for export, transporting goods (exports as well as imports) to and from ports (and manufacturing the trucks to transport them), warehousing traded goods (and manufacturing the materials used to build the warehouses), financing them, advertising them, etc. In addition, exports and imports make an economy more efficient and that efficiency in turn generates additional indirect output and related jobs. These latter impacts are likely to exceed the direct impacts. But net flows (the trade deficit) only capture a small piece of what creates output and related jobs.

Results

Impact on output of total trade

The impact of US–Canada trade on US national and state output is significant. Table 6 shows that trade with Canada boosted national output by $327 billion, or 2.6 percent of total GDP. Services sector output is heavily linked to trade with Canada. This includes not only services output related to services exports and imports, but also services output related to goods exports and imports (for example, wholesaling and retailing goods exported to or imported from Canada).

---

[4]The trade deficit is simply an accounting identity: exports less imports. What matters for evaluating the full impact of trade on any economy is not this net piece of the GDP total, but the impact of total exports and imports directly as well as indirectly on the economy as a whole.

Table 6: US National Output Related to Trade with Canada, 2005 (billions of US dollars)

Total	**$327.0**
Primary Sectors (agriculture, forestry, fishing, mining)	0.4
Construction	10.9
Manufacturing	37.8
Services	278.0
Transportation and utilities	*9.2*
Wholesale and retail trade	*52.3*
Finance and insurance	*37.7*
Other services	*178.8*
Information	*15.8*
Professional and technical	*24.1*
Management	*6.4*
Rental, leasing and real estate	*59.4*
Accommodation and food	*12.7*
Other consumer and public services	*60.3*

Source: authors' estimates

Trade with Canada also accounted for important shares of US state-level output for many states (Table 7). Not surprisingly, the states with the largest populations recorded the largest values of output related to trade with Canada: California, $43 billion (13.3 percent of the national total), Texas, $24 billion (7.3 percent of the total), New York, $26 billion (nearly 8 percent of the total), and Florida, $17 billion (5.1 percent of the total). Together these four states accounted for one-third of total national output related to trade with Canada. States for which trade with Canada accounted for relatively large shares of state output included Indiana, Delaware and Washington (2.9 percent each), and Connecticut, Iowa, Kansas, Michigan, North Carolina, Ohio and Wisconsin (2.8 percent each).

Table 7: State Output Related to Trade with Canada, 2005

(millions of US dollars)

	Value of Output	Share of Total Output	State Share of Total Output Related to Trade		Value of Output	Share of Total Output	State Share of Total Output Related to Trade
Alabama	$4,008	2.64%	1.23%	Montana	$672	2.25%	0.21%
Alaska	646	1.64	0.20	Nebraska	1,765	2.50	0.54
Arizona	5,574	2.57	1.70	Nevada	3,000	2.69	0.92
Arkansas	2,326	2.68	0.71	New Hampshire	1,496	2.72	0.46
California	43,564	2.69	13.32	New Jersey	11,737	2.72	3.59
Colorado	5.412	2.50	1.66	New Mexico	1,487	2.16	0.45
Connecticut	5,387	2.78	1.65	New York	25,697	2.68	7.86
Delaware	1,617	2.86	0.49	North Carolina	9,786	2.82	2.99
D.C.	1,848	2.26	0.57	North Dakota	566	2.32	0.17
Florida	16,946	2.52	5.18	Ohio	12,201	2.77	3.73
Georgia	9,758	2.68	2.98	Oklahoma	2,636	2.17	0.81
Hawaii	1,335	2.47	0.41	Oregon	3,956	2.74	1.21
Idaho	1,222	2.59	0.37	Pennsylvania	12,986	2.66	3.97
Illinois	15,352	2.74	4.70	Rhode Island	1,159	2.65	0.35
Indiana	6,993	2.93	2.14	South Carolina	3,811	2.72	1.17
Iowa	3,186	2.81	0.97	South Dakota	784	2.54	0.24
Kansas	2,710	2.81	0.97	Tennessee	6,142	2.66	1.88
Kentucky	3,696	2.63	1.13	Texas	23,985	2.42	7.34
Louisiana	4,036	2.40	1.23	Utah	2,330	2.57	0.71
Maine	1,152	2.56	0.35	Vermont	598	2.59	0.18
Maryland	6,240	2.53	1.91	Virginia	9,242	2.63	2.83
Massachusetts	8,722	2.68	2.67	Washington	6,943	2.90	2.12
Michigan	10,360	2.75	3.17	West Virginia	1,209	2.28	0.37
Minnesota	6,324	2.70	1.93	Wisconsin	6,047	2.80	1.85
Mississippi	2,088	2.57	0.64	Wyoming	459	1.68	0.14
Missouri	5,788	2.68	1.77	US Total	326,984	2.63	100.0

Source: authors' estimates

Impact of total trade on US jobs

As a result of the boost Canadian trade gives to US output, US jobs are also supported, both directly (in the manufacture of goods or production of services for export, for example) and indirectly (in sectors that get the goods and services out the door and

across the border to Canada). Jobs related to importing also span the service sectors, and include jobs related to transporting, wholesaling, warehousing, advertising, financing and retailing products imported from Canada, for example. In addition, it is important to note that producer services are also key inputs to manufacturing, so that goods exports indirectly support services. Our model incorporates the impact of job losses due to import competition; thus the results are net of any negative impacts of imports.

We report our estimates in Table 8. The results indicate that trade with Canada in 2005 supported 7.1 million net direct and indirect American jobs. More than half a million of these jobs are in the manufacturing sector. Services sectors account for the largest share of jobs related to trade with Canada, including such high-wage occupations as finance and insurance, legal, managerial, advertising and other professional services.

Table 8: National Employment Related to Trade with Canada, 2005

Total	**7,111,714**
Primary Sectors (agriculture, forestry, fishing, mining)	266,746
Construction	55,464
Manufacturing	522,864
Services	6,266,641
Transportation and utilities	*271,359*
Wholesale and retail trade	*235,832*
Finance and insurance	*328,202*
Other services	*5,431,247*
Information	*193,433*
Professional and technical	*391,731*
Management	*63,187*
Rental, leasing and real estate	*325,219*
Accommodation and food	*550,104*
Other consumer and public services	*3,907,573*

Source: authors' estimates, based on CGE results discussed in Appendix A

At the state level, every state experiences positive job effects from trade (exports and imports) with Canada (Table 9). The largest absolute net numbers of jobs supported by trade

with Canada were in California (832,000), Texas (522,000), New York (469,000) and Florida (405,000). Collectively, across these four states we estimate 2.2 million jobs supported by trade with Canada. For individual states, the job gains are generally in a range of 4 to 5 percent of total state-wide employment.

Table 9: State Employment Related to Trade with Canada, 2005

	Number of Jobs	Share of Total Jobs	State Share of Total Jobs Related to Trade		Number of Jobs	Share of Total Jobs	State Share of Total Jobs Related to Trade
Alabama	100,486	4.00%	1.41%	Montana	24,368	3.97%	0.34%
Alaska	19,332	4.42	0.27	Nebraska	49,697	4.09	0.86
Arizona	128,862	3.98	1.81	Nevada	61,219	4.01	0.92
Arkansas	63,323	4.07	0.89	New Hampshire	32,668	3.91	0.46
California	832,178	4.05	11.70	New Jersey	206,778	4.14	2.91
Colorado	123,794	4.03	1.74	New Mexico	44,418	4.18	0.62
Connecticut	90,192	4.15	1.27	New York	468,703	4.36	6.59
Delaware	21,332	4.04	0.30	North Carolina	208,480	4.08	2.93
D.C.	39,999	5.03	0.56	North Dakota	18,798	3.98	0.26
Florida	404,713	3.99	5.69	Ohio	276,621	4.07	3.89
Georgia	211,676	4.07	2.98	Oklahoma	81,177	3.97	1.16
Hawaii	36,893	4.42	0.52	Oregon	88,649	3.98	1.25
Idaho	33,601	3.87	0.47	Pennsylvania	295,230	4.14	4.15
Illinois	304,514	4.10	4.28	Rhode Island	25,876	4.25	0.36
Indiana	147,794	4.02	2.08	South Carolina	95,329	4.03	1.34
Iowa	77,912	3.96	1.10	South Dakota	21,426	3.98	0.30
Kansas	72,844	4.04	1.02	Tennessee	145,932	4.02	2.05
Kentucky	95,928	4.03	1.35	Texas	521,759	3.99	7.34
Louisiana	101,947	4.14	1.43	Utah	61,309	4.06	0.86
Maine	33,289	4.05	0.47	Vermont	17,410	4.11	0.24
Maryland	140,334	4.22	1.97	Virginia	197,038	4.17	2.77
Massachusetts	172,253	4.19	2.42	Washington	152,914	4.10	2.15
Michigan	221,492	4.02	3.11	West Virginia	36,925	4.06	0.52
Minnesota	141,194	4.03	1.99	Wisconsin	141,404	4.00	1.99
Mississippi	61,759	4.10	0.87	Wyoming	14,095	3.91	0.20
Missouri	144,851	4.05	2.04	US Total	7,111,714	4.08	100.0

Source: authors' estimates

172

Services sector income and employment

Close inspection of Tables 6 and 8 shows that most of our income and employment estimates are concentrated in services. There are four main reasons that, combined, lead to this outcome. The first is, quite simply, that the US economy is largely a services economy. In 2005, according to data from the US Bureau of Economic Analysis (Department of Commerce), services (including construction) accounted for 83 percent of non-farm private employment and 75 percent of private gross product. Trade with Canada means a more efficient overall US economy, and hence a general increase in economic activity, including services. Any increase in activity and employment will be largely focused on services. Second is that we are modeling direct trade in services. We capture direct linkages between exports to Canada and services production in the United States. Services are an important part of overall trade between the United States and Canada. The third reason is that manufacturing in the United States is actually quite services-intensive (Francois and Woerz 2007), so that a boost to manufacturing activity from exports to Canada has important implications for demand for intermediate services. Fourth, because we are looking at general equilibrium effects, our estimates include income linkages to services demand. This means that higher incomes lead to more demand for (and jobs linked to) consumer services. This last effect is missing from analyses that just focus on production-based input–output linkages.

Impact of changes in trade volumes and costs

Given, then, that trade supports output and jobs, it stands to reason that changes in trade and trade costs would have an impact on output and jobs. Table 10 shows the impact that a 1 percent change in trade volumes or trade costs would have on employment (or output) (referred to as "elasticities"; more detailed estimates are provided in Appendix B).

The trade-volume elasticities can be interpreted as follows. From the first row, a 1 percent increase in trade with Canada

supports a 0.038 percent increase in US employment. At the same time, a 1 percent change in trade implies that total income rises by 0.026 percent across the United States. From the values reported in Table B-1 of Appendix B, for California a 1 percent increase in trade implies a 0.037 percent increase in state employment and a 0.027 percent increase in state income.

The trade-cost elasticities in the tables in Appendix B have a similar interpretation. From Table 10, a 1 percent increase in the cost of cross-border trade between the United States and Canada implies a 0.156 percent loss in employment and a 0.103 percent loss of state income. A more detailed set of trade-cost elasticities at the state level is provided in the tables in Appendix B.

Table 10: The Marginal Impact of Changes in Trade Volumes and Costs (Percent)

	Change in total state employment	Change in total Gross State Product
1% increase in trade volumes	0.038%	0.026%
1% increase in cost of trade	-0.156%	-0.103%

Conclusion

In this paper, we have explored the impact of US trade with Canada on the pattern of employment and output across US states. This has been accomplished by using a multi-region, global CGE model to estimate the economy-wide impact of US–Canadian trade. The results of this analysis indicate that the trade relationship between the United States and Canada is a definite net plus for the United States[5]. Accordingly, policies that reduce the flow of goods and services between Canada and the United States result in adverse impacts on jobs and growth in the United States, impacts that are felt in every US state. For example, policy actions that have the effect of reducing US exports of goods or services to Canada would directly reduce US goods and services output, and related jobs. Perhaps less expected by policy-makers

---

[5]We speculate that a similar analysis for Canada would demonstrate parallel benefits to Canadian output and employment.

is the finding that actions that reduce US imports of goods and services from Canada would also have a negative net impact on US output and related employment. This should not be surprising, given the integrated nature of North American manufacturing industries and the important role of services in these continental industries. The results reported here suggest that these broader impacts should be factored into policy considerations with likely impacts on trade flows.

References

Dimaranan, B. V. and R. A. McDougall. 2002. Global trade, assistance, and production: the GTAP 5 data base. West Lafayette, Indiana: Center for Global Trade Analysis, Purdue University.

Francois, J.F. and L.M. Baughman. 2004. US–Canadian trade and US state-level production and employment. In *Trade Policy Research 2004*, ed. John M. Curtis and Dan Ciuriak. Ottawa: Department of Foreign Affairs and International Trade.

Francois, J.F. and J. Woerz. 2007. Producer services, manufacturing linkages, and trade. Discussion paper, Tinbergen Institute, Rotterdam.

Hertel, T.W., E. Ianchovichina and B.J. McDonald. 1997. Multiregion general equilibrium modeling. In *Applied Methods for Trade Policy Analysis: A Handbook*, ed. J.F. Francois and K.A. Reinert. New York: Cambridge Univ. Press.

Huff, K. R. McDougall and T. Walmsley. 2000. Contributing input–output tables to the GTAP data base. GTAP Technical Paper No. 01. West Lafayette, Indiana: GTAP Consortium.

McDonald, S. and K. Thierfelder. 2003. Deriving a global social accounting matrix from GTAP version 5 data. West Lafayette, Indiana: GTAP Consortium.

Reinert, K.A. and D.W. Roland-Holst. 1997. Social accounting matrices. In *Applied Methods for Trade Policy Analysis: A Handbook*, ed. J.F. Francois and K.A. Reinert. New York: Cambridge Univ. Press.

Appendix A: Methodology

Different options are available to estimate trade linkages to employment and output. One involves manipulation of input–output tables to map the linkages between exports and/or imports to labour demand and total output across sectors. Such an approach presents several problems, however. The first is that the shares in the base data basically fix the structure of production and demand. Second, there may be double counting, as the net effect of exports and imports is not the simple sum of export effects and import effects. Third, such an approach may also overestimate effects unless the impact of substitution toward trade with the rest of the world is also included.

To address these various issues, we applied a computable multi-sector model of the US economy. Computable general equilibrium (CGE) models are characterized by an input–output structure (based on regional and national input–output and employment tables) that explicitly link industries in a value-added chain from primary goods, over continuously higher stages of intermediate processing, to the final assembling of goods and services for consumption. Inter-sectoral linkages are direct, like the input of steel in the production of transport equipment, and indirect, via intermediate use in other sectors. The model captures these linkages by modelling firms' use of factors of production (labour and capital) and intermediate inputs. The most important aspects of the model can be summarized as follows: it covers all world trade and production, and it includes intermediate linkages between sectors.

Data

Our data come from a number of sources. Data on production and trade are based on national social-accounting data linked through trade flows (see Reinert and Roland-Holst 1997). The input–output structure of our data is drawn from the most recent version of the Global Trade Analysis Project dataset, version 6.2 (Dimaranan and McDougall 2006). In this version of the

dataset, the underlying input–output table for the United States is for the year 2004. (Earlier versions of GTAP 6 are based on 1992 input–output tables.) The GTAP version 6.2 dataset itself is benchmarked to 2001 values (where the social-accounting data have been rebalanced based on the input–output coefficients, combined with values of production, wages, and output in 2001, and also 2001 trade values). Because the data are structured to reflect the value flows in the US economy in 2001, we built a modified database that reflects the US and Canadian economies (production and trade) in 2005. Our 2005 database includes detailed national input–output flows (from the GTAP tables combined with 2005 value data), trade, and final demand structures. The basic social-accounting and trade data are supplemented with trade policy data, including additional data on tariffs and non-tariff barriers. The data are further supplemented with data from the US Department of Labor on state-level employment and from the US Bureau of Economic Analysis on state-level output. These data allow us to map nationwide effects to state-level changes in employment and output.

The data on tariffs are taken from the World Trade Organization's integrated database, with supplemental information from the World Bank's recent assessment of detailed pre- and post-Uruguay Round tariff schedules and from the UNCTAD/World Bank WITS dataset. All of this tariff information has been concorded to GTAP model sectors within the version 6.2 database. The sectors in the model are shown in Table A-1. Regions are aggregated into the United States, Canada, Mexico, and rest-of-world.

Table A-1: Model Sectors

Model Sectors	Corresponding GTAP Sectors
1. Agriculture, forestry, and fisheries	1–14
2. Mining	15, 16, 17, 18
3. Utilities	43–45
4. Construction	46
5. Durable goods manufacturing	30,34–42
6. Nondurable goods manufacturing	19–29,31-33
7. Wholesale and retail trade	47
8. Transportation	48, 49, 50
9. Information services	51
10. Finance and insurance	52–53
11. Other business services	54
12. Other consumer services	55
13. Real estate	57
14. Public services	56

The Model

Single representative, composite households comprise each region, with expenditures allocated over personal consumption and savings. The composite household owns endowments of the factors of production and receives income by selling them to firms. It also receives income from tariff revenue and rents accruing from import and export quota licenses (when applicable). Part of the income is distributed as subsidy payments to some sectors, primarily in agriculture.

On the production side, in all sectors, firms employ domestic production factors (capital, labour and land) and intermediate inputs from domestic and foreign sources to produce outputs in the most cost-efficient way that technology allows. Capital stocks are fixed at a national level. Firms are competitive, and employ capital and labour to produce goods and services subject to constant returns to scale[6]. Products from different regions are

---

[6]Compared to dynamic CGE models and models with alternative market structures, the present assumption of constant returns to scale with a fixed capital stock is closest in approach to older studies based on pure input–output modelling of trade and employment linkages. In the present context, it

178

assumed to be imperfect substitutes in accordance with the so-called Armington assumption. Table A-2 shows the trade elasticities used to model Armington demand for imports[7].

Table A-2: Substitution Elasticities

		Substitution Elasticities		
		Between competing sources of imports	Between domestic products and imports	Between capital and labour (i.e. value added)
		A	B	C
1.	Agriculture, forestry, fisheries	13.44	6.72	0.20
2.	Mining	5.60	2.80	1.26
3.	Utilities	3.80	1.90	1.40
4.	Construction	7.63	3.82	1.26
5.	Durable goods manufacturing	6.26	3.13	1.22
6.	Nondurable goods manufacturing	3.80	1.90	1.68
7.	Trade	3.80	1.90	1.68
8.	Transportation services	3.80	1.90	1.26
9.	Information services	3.80	1.90	1.26
10.	Finance and insurance	3.80	1.90	1.26
11.	Other business services	3.80	1.90	1.26
12.	Other consumer services	3.80	1.90	1.26
13.	Real estate	3.80	1.90	1.26
14.	Public services	13.44	6.72	0.20

Source: GTAP database, version 6.2

The trade substitution elasticities reported in Table A-2 show the ease with which imports can be substituted for each other (column A), and the ease with which they can be substituted for domestic goods (column B). For example, for durable goods a 1 percent increase in the price of Canadian imports

---

can be viewed as generating a lower-bound estimate of effects relative to alternative CGE modelling structures.

Model results depend on the assumptions made concerning the underlying trade elasticities. The elasticities used here are the standard set of elasticities for the Global Trade Analysis Project database. We performed a sensitivity analysis to show the impact of varying these assumptions; the results are shown in Table A-3.

causes a 7.63 percent increase in the ratio of imports of non-Canadian to Canadian-source imports. Similarly, a 1 percent increase in the price of imports of durable goods leads to a 3.82 percent increase in the ratio of domestic to imported consumption. In other words, elasticities quantify the degree to which firms and consumers shift between imports and domestic goods as relative prices change.

We were interested in the impact of trade with Canada on state economies given the current US wage structure. To quantify these linkages, we employ a labour-market closure (equilibrium condition); under this approach, we fix wages at current levels and force employment levels to adjust. This provides a direct estimate of the jobs supported, at current wage levels, by the current level of trade.

Experiments

The experiments conducted with the model involve imposing changes in US–Canada trade. This allows us to trace changes at the border as they work through the US economy. Our experiment involved one change to 2005 trade flows: a shut-down of Canadian goods and services exports to the United States simultaneously with a shutdown in US goods and services exports to Canada[8].

---

[8]This is accomplished by making a set of bilateral export taxes endogenous, while making trade quantities exogenous and then reducing them by target amounts, which is appropriate since the relevant question is the benefit of current conditions of trade.

Table A-3: Sensitivity Analysis with Respect to Trade Elasticities: Ranging One Standard Deviation above and below Mean Values

	state employment			gross state product		
	lower	mean	upper	lower	mean	upper
United States	2,948,103	7,111,714	11,275,326	135,548	326,984	518,419
Alabama	41,631	100,486	159,340	1,660	4,008	6,355
Alaska	7,965	19,332	30,698	266	646	1,026
Arizona	53,442	128,862	204,282	2,311	5,574	8,836
Arkansas	26,057	63,323	100,590	957	2,326	3,694
California	345,296	832,178	1,319,060	18,076	43,564	69,052
Colorado	51,275	123,794	196,314	2,242	5,412	8,582
Connecticut	37,558	90,192	142,826	2,243	5,387	8,531
Delaware	8,890	21,332	33,774	674	1,617	2,560
District of Columbia	16,832	39,999	63,165	778	1,848	2,919
Florida	168,169	404,713	641,256	7,042	16,946	26,851
Georgia	87,758	211,676	335,593	4,045	9,758	15,470
Hawaii	15,418	36,893	58,368	558	1,335	2,112
Idaho	13,845	33,601	53,356	503	1,222	1,940
Illinois	126,169	304,514	482,859	6,361	15,352	24,343
Indiana	61,058	147,794	234,530	2,889	6,993	11,096
Iowa	32,140	77,912	123,685	1,314	3,186	5,058
Kansas	30,021	72,844	115,668	1,117	2,710	4,304
Kentucky	39,411	95,928	152,445	1,518	3,696	5,874
Louisiana	42,078	101,947	161,817	1,666	4,036	6,406
Maine	13,868	33,289	52,709	480	1,152	1,825
Maryland	58,576	140,334	222,093	2,605	6,240	9,876
Massachusetts	71,799	172,253	272,707	3,636	8,722	13,808
Michigan	91,792	221,492	351,193	4,293	10,360	16,427
Minnesota	58,496	141,194	223,892	2,620	6,324	10,028
Mississippi	25,523	61,759	97,996	863	2,088	3,313
Missouri	59,997	144,851	229,704	2,397	5,788	9,178
Montana	10,054	24,368	38,682	277	672	1,067
Nebraska	20,494	49,697	78,900	728	1,765	2,802
Nevada	25,377	61,219	97,060	1,244	3,000	4,757
New Hampshire	13,600	32,668	51,736	623	1,496	2,370
New Jersey	85,993	206,778	327,564	4,881	11,737	18,592
New Mexico	18,382	44,418	70,453	616	1,487	2,359
New York	195,849	468,703	741,556	10,738	25,697	40,657
North Carolina	86,632	208,480	330,329	4,067	9,786	15,506
North Dakota	7,735	18,798	29,860	233	566	900
Ohio	114,639	276,621	438,604	5,056	12,201	19,345
Oklahoma	33,569	82,177	130,785	1,077	2,636	4,195
Oregon	36,649	88,649	140,649	1,635	3,956	6,276
Penn- sylvania	122,627	295,230	467,832	5,394	12,986	20,578
Rhode Island	10,807	25,876	40,944	484	1,159	1,835
South Carolina	39,642	95,329	151,015	1,585	3,811	6,036
South Dakota	8,865	21,426	33,987	324	784	1,243
Tennessee	60,248	145,932	231,616	2,536	6,142	9,749
Texas	214,502	521,759	829,016	9,860	23,985	38,109
Utah	25,392	61,309	97,226	965	2,330	3,696
Vermont	7,248	17,410	27,572	249	598	948
Virginia	81,904	197,038	312,172	3,842	9,242	14,643
Washington	63,483	152,914	242,345	2,882	6,943	11,003
West Virginia	15,175	36,925	58,676	497	1,209	1,921
Wisconsin	58,465	141,404	224,344	2,500	6,047	9,594
Wyoming	5,670	14,095	22,521	185	459	733

Based on Gaussian quadrature, where standard error=0.5*elasticity and where nesting is imposed so that the lower-level Armington elasticity = ½ the upper-level Armington elasticity. Given the actual uncertainty surrounding GTAP trade elasticities, we have overestimated the confidence bounds.

Appendix B: Marginal Impact of Trade on Jobs and GSP: Elasticities Analysis

The first two columns in Table B-1 at the end of this Appendix B provide estimates of the marginal impact of trade on state-level employment and gross state product (GSP). Technically, the numbers in the table are elasticities. This means that they measure the percentage change in employment (or GSP) associated with a 1 percent change in trade. The output and employment elasticities are defined as follows:

$$\text{Output elasticity} = \%\Delta GSP\ /\ \%\Delta Trade \qquad (1)$$

$$\text{Employment elasticity} = \%\Delta Employment\ /\ \%\Delta Trade \qquad (2)$$

Columns 2 and 3 in Table B1 report the results. For total US GSP across all states, the output elasticity is 0.026 (top row, column 3). This means that a 10 percent drop in trade maps to a 10%*0.026=0.26 percent drop in total state GSP. For a full closure of trade, the value is approximately 100%*0.026=2.6%. From Table 6 of the main text, our exact estimate is 2.63%, approximated by the value implied by the elasticity in the Table B-1. The same relationships hold with all values in Tables B-1, B-2 and B-3.

Similarly, for total US employment across all states, the employment elasticity is 0.038 (top row, column 2). This implies that a 10 percent increase (or reduction) in trade would imply a 0.38 percent increase (or drop) in US jobs.

The state-level impacts are interpreted in a similar fashion. Thus, for California, a 1 percent increase in trade implies a 0.037 percent increase in state employment and a 0.027 percent increase in state income.

Columns 4 and 5 of Table B-1 focus on a different aspect of the same issue. They also report elasticities. However, these involve changes in state employment and GSP that result from a 1 percent increase in trade costs (the cost of delivering goods

across the border, measured as a share of the price of goods and services traded.) From the table, a 1 percent increase in the cost of trade implies a 0.156 percent drop in US employment and a 0.103 percent drop in incomes at the state level. Choosing California again as an example, this means that an increase in border costs equal to 1 percent (10 percent) of the price of traded goods and services implies a 0.155 percent (1.55 percent) drop in state employment in California and a 0.104 percent (1.04 percent) drop in state income.

Table B-2 provides a state breakdown for employment by broad sectors. The values in the table are again elasticities. However, this time they are reported for primary, manufacturing, and service sector employment linked to the level of trade with Canada.

Finally, Table B-3 provides a breakdown for state-level GSP by broad sectors.

Table B-1: State Employment and GSP Elasticities: Trade with Canada

	% impact of 1% change in trade volumes		% impact of 1% change in cost of trade	
	jobs	GSP	jobs	GSP
United States	0.038	0.026	-0.156	-0.103
Alabama	0.037	0.026	-0.154	-0.103
Alaska	0.041	0.017	-0.170	-0.062
Arizona	0.037	0.026	-0.152	-0.099
Arkansas	0.038	0.026	-0.156	-0.105
California	0.037	0.027	-0.155	-0.104
Colorado	0.037	0.026	-0.154	-0.096
Connecticut	0.038	0.028	-0.159	-0.109
Delaware	0.037	0.028	-0.155	-0.113
District of Columbia	0.047	0.024	-0.194	-0.086
Florida	0.037	0.026	-0.153	-0.097
Georgia	0.038	0.027	-0.156	-0.106
Hawaii	0.041	0.026	-0.169	-0.095
Idaho	0.036	0.026	-0.148	-0.099
Illinois	0.038	0.027	-0.157	-0.107
Indiana	0.037	0.028	-0.154	-0.115
Iowa	0.037	0.027	-0.151	-0.110
Kansas	0.038	0.026	-0.155	-0.099
Kentucky	0.038	0.026	-0.154	-0.103
Louisiana	0.038	0.024	-0.159	-0.096
Maine	0.038	0.026	-0.155	-0.100
Maryland	0.039	0.026	-0.162	-0.099
Massachusetts	0.039	0.027	-0.161	-0.104
Michigan	0.037	0.027	-0.154	-0.106
Minnesota	0.038	0.027	-0.154	-0.105
Mississippi	0.038	0.026	-0.157	-0.100
Missouri	0.038	0.027	-0.155	-0.105
Montana	0.037	0.023	-0.152	-0.086
Nebraska	0.038	0.025	-0.156	-0.097
Nevada	0.036	0.028	-0.153	-0.103
New Hampshire	0.036	0.027	-0.150	-0.106
New Jersey	0.038	0.027	-0.159	-0.108
New Mexico	0.039	0.022	-0.160	-0.082
New York	0.040	0.027	-0.167	-0.105
North Carolina	0.038	0.028	-0.157	-0.113
North Dakota	0.037	0.023	-0.152	-0.089
Ohio	0.038	0.027	-0.156	-0.108
Oklahoma	0.037	0.022	-0.152	-0.084
Oregon	0.037	0.027	-0.152	-0.105
Penn- sylvania	0.038	0.026	-0.159	-0.105
Rhode Island	0.039	0.026	-0.163	-0.103
South Carolina	0.037	0.027	-0.155	-0.107
South Dakota	0.037	0.025	-0.152	-0.097
Tennessee	0.037	0.026	-0.154	-0.105
Texas	0.037	0.024	-0.152	-0.095
Utah	0.038	0.026	-0.155	-0.099
Vermont	0.038	0.026	-0.157	-0.100
Virginia	0.038	0.027	-0.160	-0.103
Washington	0.038	0.026	-0.157	-0.100
West Virginia	0.038	0.023	-0.156	-0.089
Wisconsin	0.037	0.027	-0.153	-0.110
Wyoming	0.036	0.018	-0.149	-0.064

Note: all elasticities are significant at the 5% level, based on Gaussian quadrature
sensitivity analysis of estimates with respect to uncertainty about values of trade elasticities.

184

Table B-2: Detailed State Employment Elasticities: Trade with Canada

	% impact of a 1% change in trade volumes on employment		
	primary	manufacturing	services
United States	0.043	0.034	0.039
Alabama	0.050	0.034	0.038
Alaska	0.011	0.038	0.043
Arizona	0.028	0.032	0.037
Arkansas	0.061	0.035	0.040
California	0.036	0.034	0.038
Colorado	0.029	0.033	0.038
Connecticut	0.054	0.033	0.039
Delaware	0.066	0.036	0.038
District of Columbia	0.019	0.037	0.047
Florida	0.034	0.033	0.037
Georgia	0.054	0.035	0.038
Hawaii	0.069	0.037	0.041
Idaho	0.058	0.034	0.038
Illinois	0.070	0.034	0.039
Indiana	0.091	0.033	0.039
Iowa	0.155	0.034	0.039
Kansas	0.059	0.034	0.040
Kentucky	0.059	0.034	0.040
Louisiana	0.015	0.035	0.040
Maine	0.031	0.035	0.039
Maryland	0.047	0.035	0.039
Massachusetts	0.028	0.033	0.039
Michigan	0.060	0.032	0.038
Minnesota	0.109	0.034	0.039
Mississippi	0.052	0.033	0.040
Missouri	0.128	0.034	0.039
Montana	0.048	0.034	0.039
Nebraska	0.125	0.035	0.040
Nevada	0.009	0.033	0.037
New Hampshire	0.038	0.032	0.037
New Jersey	0.050	0.036	0.039
New Mexico	0.022	0.033	0.040
New York	0.047	0.034	0.041
North Carolina	0.061	0.035	0.039
North Dakota	0.088	0.033	0.040
Ohio	0.066	0.033	0.039
Oklahoma	0.029	0.033	0.040
Oregon	0.053	0.033	0.039
Pennsylvania	0.043	0.034	0.039
Rhode Island	0.025	0.033	0.040
South Carolina	0.057	0.035	0.038
South Dakota	0.147	0.033	0.039
Tennessee	0.130	0.034	0.039
Texas	0.022	0.034	0.038
Utah	0.035	0.033	0.039
Vermont	0.056	0.033	0.039
Virginia	0.056	0.034	0.039
Washington	0.044	0.033	0.039
West Virginia	0.015	0.034	0.040
Wisconsin	0.119	0.034	0.039
Wyoming	0.010	0.034	0.040

Table B-3: Detailed GSP Elasticities: Trade with Canada

	% impact of a 1% change in trade volumes on GSP		
	primary	manufacturing	services
United States	0.004	0.020	0.025
Alabama	0.005	0.020	0.024
Alaska	0.000	0.020	0.023
Arizona	0.005	0.020	0.025
Arkansas	0.007	0.020	0.024
California	0.007	0.020	0.026
Colorado	0.002	0.020	0.026
Connecticut	0.008	0.020	0.026
Delaware	0.010	0.020	0.027
District of Columbia	0.005	0.020	0.023
Florida	0.009	0.020	0.025
Georgia	0.007	0.020	0.025
Hawaii	0.009	0.020	0.025
Idaho	0.009	0.020	0.025
Illinois	0.006	0.020	0.025
Indiana	0.006	0.020	0.024
Iowa	0.009	0.020	0.024
Kansas	0.006	0.020	0.024
Kentucky	0.004	0.020	0.023
Louisiana	0.001	0.020	0.024
Maine	0.010	0.020	0.024
Maryland	0.007	0.020	0.025
Massachusetts	0.008	0.020	0.025
Michigan	0.007	0.020	0.025
Minnesota	0.008	0.020	0.025
Mississippi	0.005	0.020	0.024
Missouri	0.007	0.020	0.025
Montana	0.005	0.020	0.024
Nebraska	0.010	0.020	0.024
Nevada	0.001	0.020	0.028
New Hampshire	0.008	0.020	0.026
New Jersey	0.007	0.020	0.026
New Mexico	0.001	0.020	0.024
New York	0.007	0.020	0.026
North Carolina	0.009	0.020	0.025
North Dakota	0.006	0.020	0.024
Ohio	0.005	0.020	0.024
Oklahoma	0.001	0.020	0.024
Oregon	0.010	0.020	0.025
Pennsylvania	0.004	0.020	0.024
Rhode Island	0.008	0.020	0.025
South Carolina	0.008	0.020	0.024
South Dakota	0.009	0.020	0.025
Tennessee	0.007	0.020	0.024
Texas	0.001	0.020	0.025
Utah	0.002	0.020	0.025
Vermont	0.008	0.020	0.024
Virginia	0.005	0.020	0.025
Washington	0.010	0.020	0.025
West Virginia	0.001	0.020	0.023
Wisconsin	0.009	0.020	0.024
Wyoming	0.000	0.020	0.025

The Canadian Automotive Market

Johannes Van Biesebroeck[*]

Executive summary

The automotive sector is Canada's largest manufacturing sector, accounting for 12% of its manufacturing GDP and 25% of its manufacturing trade. The principal objective of this study is to calculate the impact of changes in Canada's trade policy on the automotive sector. The study is organized in five sections: the first identifies current and future trends in the industry; the second contains an econometric model to analyse the market effects of four trade policy scenarios on automobile production; the third identifies the impact of trade policy on foreign direct investment; the fourth contains an analysis of the market effects of trade policy changes on the aftermarket auto parts sector; and the last section of the study discusses the future direction of the automotive industry.

Current and Future Trends in the Industry

Despite record sales in North America over the past few years, the long-term trend for the automotive industry is weighted towards higher growth rates in less developed economies, particularly China, Korea, Mexico, Brazil, India and Thailand. While global production increased by a factor of six between 1950 and 2004, combined production in Canada and the United States less than doubled over the same time period. Even though Canadian exports of finished vehicles remain very strong, a

concern is the reliance on the U.S. market. From a policy perspective, there is little Canada can do about this. The export potential for vehicles produced in Canada is effectively driven by the type of vehicles foreign-owned manufacturers decide to produce in their Canadian assembly plants.

The larger growth area for the Canadian automotive industry in recent decades has been in parts and components which, by 2002, had reached 66% of total automotive employment, up from 55% in 1991. Exports of automotive parts, while also very concentrated on the United States, are slightly more diversified than is the case for vehicles.

Market Analysis: Automobile Production

The Model

The econometric analysis of the impact of trade policy on the vehicle assembly sector was conducted in three steps. First, a nested logit model was used to estimate demand at the vehicle level based on seven nests. This model selection results in higher elasticities of substitution between models in the same segment than across segments. Second, the demand model was used to calculate a number of quantities that influence the effect of policy changes including: (i) own and cross-price elasticities for each model with respect to all other models in the market; (ii) unobserved vehicle quality, from the point of view of the consumer; and (iii) the marginal costs for each vehicle that are consistent with the estimated price elasticities of demand and the observed prices. To calculate the elasticities and marginal costs, it is assumed that firms are playing a Bertrand price-setting game (i.e., a specific form of game theory) in differentiated products. Third, using the estimated demand parameters, price-elasticities and marginal costs, simulations of market equilibrium are conducted to examine the impact of elimination of Canada's 6.1% import tariff on non-NAFTA vehicles.

There are four trade policy changes simulated using this model: (i) an FTA with South Korea; (ii) an FTA with Japan; (iii) an FTA with the European Union (E.U.); and (iv) unilateral

188

abolition of the Canadian tariff on imported vehicles. An FTA is assumed to include the elimination of tariffs on imports from the partner country.

An FTA with Korea: The results of the model's application to elimination of tariffs with Korea is a decrease in average prices, an increase in average mark-ups for Korean firms and a slight decrease for foreign firms, and an increase in aggregate vehicle sales. In the end, Korean imports are estimated to increase by 9.72%, while all other foreign suppliers lose. As well, production in Canada declines by 0.53%.

An FTA with Japan: While the analysis of an FTA with Japan is similar to that of an FTA with Korea, one notable difference is that, due to compositional effects, i.e., sales shifting upmarket as prices decrease, the average sales-weighted Japanese price ends up higher with an FTA. Another is that the largest effect of this FTA would be a 3.14% decrease in imports from the E.U. because they compete with Japan-made cars in all luxury segments. In the end, Japanese imports are estimated to increase by 15.11%, while production in Canada falls by 0.94%.

An FTA with the E.U.: Due to the higher demand elasticities of the median European car in every segment, an FTA with the E.U. brings even stronger compositional effects than an FTA with Japan. In this scenario, the average price is estimated to increase as the generally expensive European vehicles gain market share. The increase in imports from the E.U. is estimated at 28.32%, while Canadian production is estimated to decrease by 0.74%.

Unilateral Tariff Elimination: Under unilateral tariff elimination by Canada, Canadian production is estimated to decline by 8,668 units annually (2.16%). While this is not nearly enough to noticeably impact assembly plant capacity decisions, no doubt employment would be affected, including in supplier plants, and some workers would face transition costs. In addition, while Korea, Japan and the E.U. all benefit under this scenario, the

189

import gains go disproportionately to the E.U., which sees its imports increase by almost 24.53% versus only 7.68% for Korea.

As demonstrated in the following table, there are increases in consumer surplus that would accrue in each of the above scenarios. However, overall domestic welfare is estimated to decrease marginally in each case. This is mainly due to the large decreases in government tariff revenues.

	FTA with:			Unilateral Tariff Elimination
	Korea	Japan	E.U.	
Aggregate effects on:				
Price (average)	-0.35%	-0.27%	0.95%	0.30%
Demand	0.25%	0.53%	0.45%	1.22%
Canadian production[1]	-0.53%	-0.94%	-0.74%	-2.16%
Imports	0.52%	1.04%	0.86%	2.37%
Consumer surplus	0.28%	0.60%	0.51%	1.37%
Tariff revenue	-21.83%	-44.84%	-36.62%	-100.00%
Domestic welfare	-0.04%	-0.04%	-0.02%	-0.08%

[1.] refers only to Canadian production of vehicles sold in Canada

Foreign Direct Investment

While a tariff on final vehicle imports provides incentives for foreign firms to establish local production capacity to avoid the tariff, current tariff levels are sufficiently low and the overcapacity in the market sufficiently large such that no significant investment impact would be expected from any of the scenarios analysed in section two. In addition, the probability that any firm will expand assembly capacity in North America beyond the currently announced plans is relatively small. In terms of potential expansion of Canadian exports of finished vehicles, this is also small. Large export volumes from Canada to the rest of the world also seems an unlikely proposition, in part due to likely increases in exports from low wage countries, only a marginal phenomenon for the moment.

Market Analysis: Aftermarket Auto Parts

In order to assess the impact of trade policy changes on the more diverse parts and components sector, a number of methodologies are used to estimate demand and supply elasticities. Simulations are then conducted to examine the impact on Canadian exports in the event of FTAs with China, South Korea and the E.U. The estimated changes in Canadian exports of automotive parts range from 10.4% to 22.2% for an FTA with China; 8.4% to 11.6% for an FTA with South Korea; and 3.4% to 7.9% for an FTA with the E.U. in view of the fact that current trade protection for the parts sector in Canada is very low. If giving up the limited protection that exists would result in lower overseas trade barriers (which tend to be higher), the net effect would likely be positive.

Future Directions and Concluding Comments

There are many factors that are likely to affect the future direction of the automotive industry in Canada, including: the types of fuels that cars will be using; whether current trends towards keeping manufacturing close to the location of the final customer remain constant; future sales volumes in North America; the location of research and development; and government policy. However, among the limited areas where government intervention may have an effect on the automotive sector, intervention in the area of trade policy is likely to have a more limited net effect on welfare than alternatives such as investment, research and development, and infrastructure support.

The study concludes that changes to Canada's trade policy would have a minimal net impact on Canada as a whole. In particular, while elimination of Canada's automotive tariff may have a modest impact on Canadian production, these losses are expected to be offset by consumer gains.

Table of Contents

1. Current and future trends

This section analyzes and documents current trends in global production and trade in the auto industry, examines Canada's position (competitiveness and technology leadership) in the global and North American auto markets, and identifies emerging trends and issues in the industry that require policy-makers' attention.

1.1 Canada's automobile industry

1.1.1 Current situation

Even though the automotive industry in North America is going from one bumper sales year to the next, the long-term trend for the industry is weighted towards higher growth rates in less developed economies. Figure 1A plots the cumulative global production of cars and light trucks, split by region. While global production increased by a factor of six between 1950 and 2004, combined production in Canada and the United States less than doubled over the same time period: average production between 1990 and 2000 was approximately 50% higher than between 1950 and 1960. In 2004, the last year of data in Figure 1.1A, production in Canada and the United States stood at 14.7 million vehicles, approximately the same as the average for the latter half of the 1990s and approximately equal to the combined output of all non-traditional producers (the rest of the world minus North America, Europe, and Japan).

The same production statistics for each region are plotted in Figure 1.1B as a fraction of global output. The declining relative importance of Canada and the United States is put in stark perspective. While these two countries accounted for almost 80% of world output in 1950, this declined to 24% in 1980. The subsequent establishment of North American assembly plants by foreign producers stabilized, even increased slightly for a while, the North American share of world output, which currently stands at 23%[1].

---

[1] Note that the higher average value of vehicles produced in North America gives the region a higher relative weight in value terms.

Figure 1.1A: Light vehicle production by region (million vehicles)
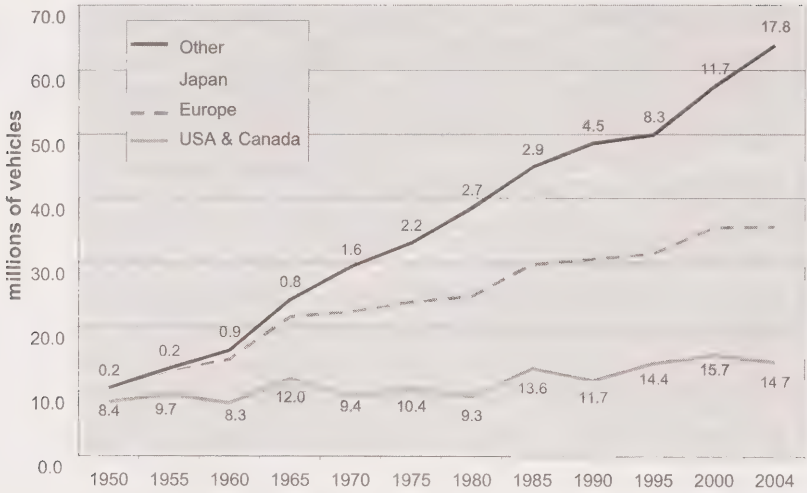
Source: Ward's World Motor Vehicle Data Book (2005)

Figure 1.1B: Light vehicle production by region (fraction of worldwide total)
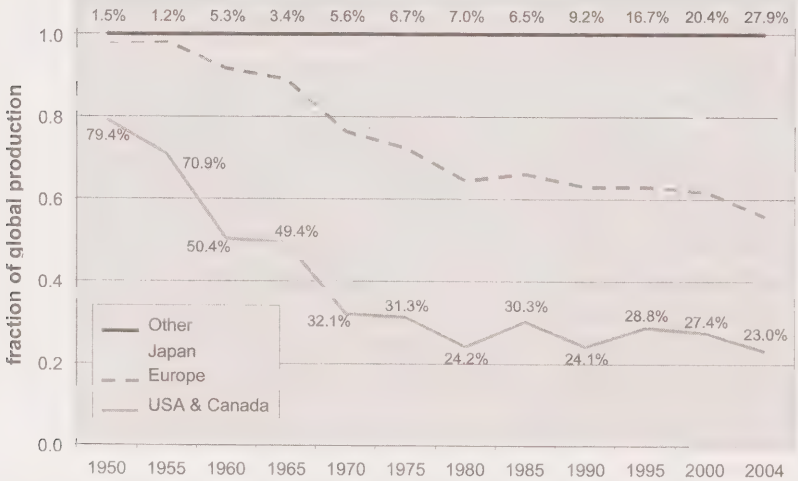
Source: Ward's World Motor Vehicle Data Book (2005)

Figure 1.1C: Light vehicle production by non-traditional producers (countries)
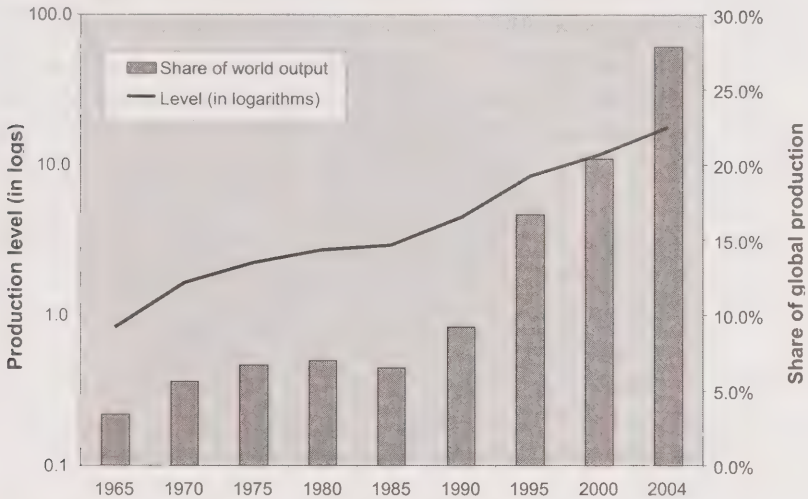
In the first 15 years after World War II, the fastest growth took place in Europe, which quickly doubled its share in world production. Subsequently, its relative importance declined somewhat, but the decline was cushioned by the more recent rise of Eastern Europe as a lower-cost manufacturing base.

Over the next 20 years, from 1965 to 1985, Japan increased its share of world production of light vehicles from 3% to 29%. In contrast with the North American and European experience, Japan's production increase was largely export driven. Import tariffs and quotas in Europe and voluntary export restrictions in the United States led Japanese producers to open up assembly plants in all their major export markets, lowering the share of Japan in world production to approximately 17% recently. The spectacular appreciation of the yen was an additional incentive for Japanese firms to establish production capacity overseas.

Finally, in the last seventeen years, most of the output growth was in non-traditional car producing countries. While the first three regions saw their combined output fluctuate between 42 and 46 million vehicles with no noticeable trend, production in the rest of the world increased from 2.9 million

197

vehicles in 1985 to 17.8m in 2004. This represents a six-fold increase in output over 19 years or a sustained annual growth rate of more than 10%. As a result, these countries produced more than one quarter of all vehicles worldwide in 2004, and this fraction has increased further in the last two years. To illustrate the importance of the output increase in these countries, Figure 1.1C plots their production level in logarithms (left scale) and their share in world output (right scale) over the last 39 years. Output growth in these countries has been remarkably constant at a very high level. All indications are that this trend will continue in the near future.

The composition of the group of "other countries" is illustrated in Figure 1.2, where the percentages indicate the share of production of each country in the group in 2004. The six most important producers are China, South Korea, Mexico, Brazil, India, and Thailand. In the last 3 years, the importance of China, and to lesser extent India, has increased further. Smaller producers are all countries producing less than 350,000 vehicles per year. One notable fact is that net exporters dominate the group of "other countries". All of the six countries depicted have higher production than domestic sales. Among the smaller producers, only Argentina is an important net exporter. Furthermore, with the exception of Brazil, the largest producers are also the countries with the fastest output growth. Given the high scale economies in vehicle production, it is no surprise that production is relatively concentrated even among emerging countries.

Focusing on the North American market, total sales in 2002 stood just under 20 million vehicles, at 19,487,556. The origin of the vehicles is depicted in Figure 1.3. While the United States accounts for almost 87% of North American sales, only 61% of vehicles are produced in that country. Foreign imports are the second most important source and Canadian production is slightly lower. Mexican production is the least important source, accounting for less than 8% of North American sales, but is growing rapidly. Almost half of all vehicles assembled in Mexico are exported and this fraction is increasing.

198

Figure 1.2: Composition of the group of 'other countries' (fraction in the group's output in 2004 is indicated)
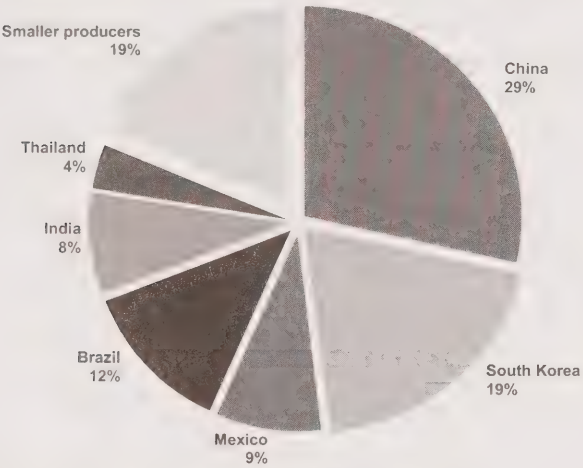
Source: Ward's World Motor Vehicle Data Book (2005)

Figure 1.3: Origin of vehicles sold in North America (2004)
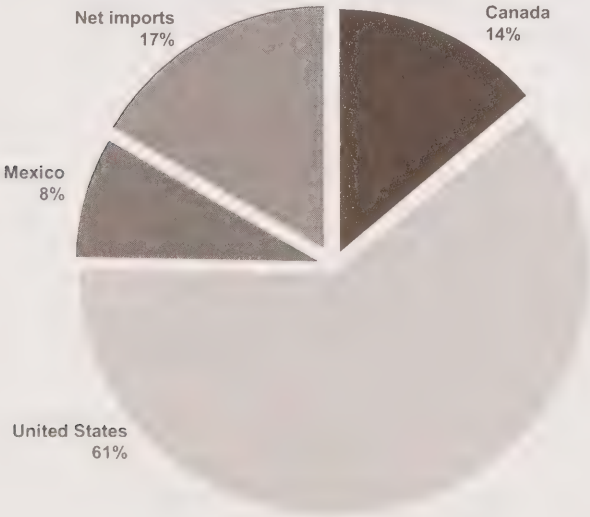
Source: Ward's Automotive Yearbook (2005)

The evolution of international trade on the North American content is also instructive. As a region, North America is running a trade deficit in vehicles that is relatively stable over time. In Figure 1.4, this is depicted by the white line. From 1972 to 2002, imports of finished vehicles fluctuated between 4.6 million units (in 1986) to a low of 1.1 million (in 1995). Net imports as a fraction of sales fluctuated between 10% and 20% in most years. North American exports are relatively unimportant and fluctuations in imports determine the trade balance almost completely. Imports started to decline in the mid-1980s when foreign producers opened their first assembly plants on the continent. In 1982, the year Honda opened its first U.S. plant, 27% of all vehicles sold in North America were imported. This declined to a mere 6.9% in 1995, after which it started to increase again, in line with the rising U.S. trade deficit for the entire economy.

The pattern for the continent as a whole is driven by the United States, which runs an even larger deficit than the region. In Figure 1.4, the blue line for the United States lies everywhere below the white line for North America. The mirror image is trade surpluses by Canada and Mexico. After the establishment of the Autopact in 1965, the Canadian industry integrated completely with the U.S. industry. Between 1972 and 1988, the trade surplus fluctuated between 7% and 43% of Canadian sales, which corresponds to an average net export of 320,000 vehicles, the vast majority to the United States. After the establishment of the Free Trade Agreement with the United States (in 1988) Canadian exports surged, even surpassing domestic sales in 1995 with the expansions of the Honda and Toyota plants coming on steam. Since then, Canadian exports have returned to normal levels, which are still 50% of domestic sales or almost 1 million vehicles.

Closures of assembly facilities in Bromont by Hyundai (1993), in Halifax by Volvo (1998), in Ste. Therese by General Motors (2002), the Pilette Road plant in Windsor by Daimler-Chrysler (2003), and the Ontario Truck plant in Oakville by Ford (2004) reduced Canadian production subsequently. The recently announced closure of the Oshawa 2 plant (2007) and

the elimination of the third shift in the Oshawa 1 plant (2006), both by General Motors, will reduce production capacity more than the newly announced assembly plant that Toyota will build in Woodstock (2007). However, it should be noted that increased production at existing plants is likely to lead to stable production levels, as forecasted by CSM.

Figure 1.4: Net trade in light vehicles for North America and individual countries — (production-sales)/sales

Mexico has done particularly well, even before the North American Free Trade Agreement took effect in 1996. Its exports grew from less than half a million vehicles when NAFTA was negotiated to over one million vehicles by 2000. This export growth is in sharp contrast with domestic sales which collapsed in 1995-96 (hence the sharp increase in Figure 4), but which have returned to the trend growth path since. Mexican sales increased by 5% per year on average over the last 25 years, only slowing to 3.6% in the last 10 years.

While the Canadian automotive industry has performed well in terms of final vehicle production, the growth rate in components has been even more remarkable. This shows up

most clearly in the employment figures for the final assembly sector versus the parts and components sector. Table 1.1 has the employment numbers for the two sub-sectors in 1991 and 2002. The employment share of parts and components in the automotive total grew from 55% in 1991 to 66% in 2002. In Section 1.2, on the vertical organization of the industry, we will discuss the parts sector in greater detail.

Table 1.1: Manufacturing employment in the Canadian automobile industry

	Employment	
	1991	2002
Vehicle assembly	53,300	51,000
Parts and Components	65,400	98,100
Source: Industry Canada		

Thus far we have discussed the Canadian automobile industry by itself and in relation to the rest of the worldwide automotive industry. It is worthwhile to stress its importance for Canadian manufacturing. Industry Canada estimates that the entire industry employs more than half a million employees in Canada: 171,002 people in automotive assembly and component manufacturing, and another 333,529 in distribution and aftermarket sales and service. Manufacturing is clustered in central Canada, in the heart of the North American auto industry, while distribution is spread across the country. It is Canada's largest manufacturing sector, accounting for 12% of the sector's GDP and 25% of manufacturing trade[2]. In 2003, Canada had an overall automotive trade surplus of $4.6 billion on flows totalling $159.1 billion. Total industry shipments stood at $69.3 billion in vehicles and $31.4 billion in parts in 2003. Production, especially in the final assembly sector, but to a lesser extent also in parts, is concentrated in South-western Ontario.

---

[2] Industry Canada, *Canada's Automotive Industry 2004*,
http: strategis ic gc ca epic internet inauto -auto.nsf en am01 22c.html

1.1.2 Future outlook

Even though Canadian exports of finished vehicles are very strong, a concern is the reliance on the U.S. market. Figure 1.5 illustrates that the vast majority of Canadian vehicle exports (HS code 8703) are destined for the United States. The graph on the left illustrates how dominant the U.S. market is, accounting for more than 98% of Canadian exports. Imports, on the other hands, are less concentrated as Japanese, Korean, and Mexican imports have grown from 26% in 1998 to almost 40% in 2004. The graph on the right illustrates the same export numbers, normalizing the 1998 levels to 1. Exports to other countries, the white line, increased noticeably, although from a very low base.

From a policy perspective, there is little Canada can do about this. The export potential for vehicles produced in Canada is entirely driven by the type of vehicles the (foreign-owned) producers decide to allocate to their Canadian assembly plants. In this respect, it is very encouraging that several Canadian plants have received the world mandate for the vehicle(s) they are assembling, meaning that no other plant produces the same vehicle.

Exports of automotive parts (HS code 8708), while also very concentrated on the United States, tend to be slightly more diversified. Throughout the 1998-2004 period, the share of Canadian parts exports going to the United States was constant around 91%. Total parts exports grew substantially over this period and exports to other countries outpaced U.S.-bound exports (see the right graph in Figure 1.6).

In 2004, exports of parts to non-U.S. destinations were worth US$896 million, almost twice the value of vehicle exports outside the United States. Moreover, while Canada had a large and growing trade deficit in vehicles with the rest of the world (excluding the United States), its trade deficit in parts declined from US$1,090m in 1998 to US$800m in 2003, although it recently jumped back up to $1,028m (in 2004).

Even for parts, the United States is by far the most important partner and the concentration of Canadian exports is increasing over time. Table 1.2 indicates that, even though the share of parts

exported to the U.S. declined only marginally from 92.5% in 1993 to 91.6% in 2004, the share of the other important export destinations increased noticeably. The five most important export destinations now account for 99.2% of Canadian parts exports. The increase is most visible for exports going to other countries (excluding the U.S.). Compared with the production statistics in Section 1.1.1, Canadian parts exports are clearly more concentrated than worldwide production. In particular, exports to Europe and Japan are much lower than expected.

The ongoing FTA negotiations with Korea could result in a more favourable import regime for Canadian parts and vehicles and increase Canadian exports to that part of the world. The recently started trade talks with Japan would work towards the same goal in Japan and the Free Trade Area of the Americas could increase Canadian exports to Latin America as well. Given the low level of trade protection, at least in terms of import tariffs, only moderate effects are expected from these initiatives. The concentration in the industry means that individual firm decisions are likely to determine the trade flows and balances. For example, as long as GM was exporting body panels to its Buick plant in China, Canada was running a large trade surplus with China. The end of these exports in 2004 instantly almost halved Canada's exports to China[3]. In terms of trade policies, all changes are in the right direction, we just expect them to be of second-order importance.

---

[3] The Asia Pacific Foundation of Canada, (2005) "The East Asian Automobile Industry: Opportunity or Threat?"

**Figure 1.5: Canadian exports of finished vehicles
(billion US$ left, 1998=1 right)**

Billion US$ 1998 = 1
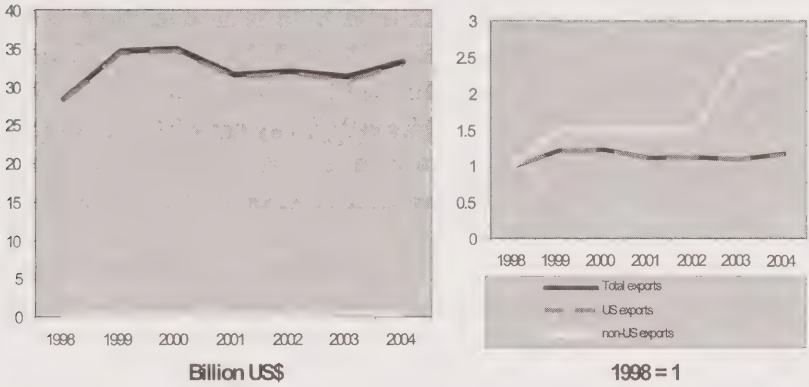
Source: U.N. Comtrade data set (online)

**Figure 1.6: Canadian exports of automotive parts and components
(US$ left, 1998=1 right)**

Billion US$ 1998 = 1
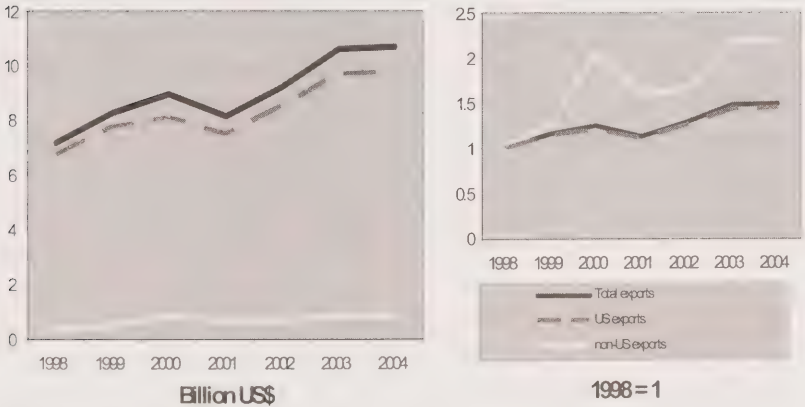
Source: U.N. Comtrade data set (online)

Table 1.2: Origin and destination of trade in parts and components (fraction of total)

	Imports			Exports	
	1993	**2003**		**1993**	**2003**
USA	0.857	0.875	USA	0.925	0.916
Share of trade excluding the U.S.			**Share of trade excluding the U.S.**		
Japan	0.451	0.427	China	0.196	0.562
Mexico	0.300	0.257	Mexico	0.208	0.213
E.U.	0.177	0.165	E.U.	0.119	0.092
China	0.002	0.054	Japan	0.085	0.039
Korea	0.033	0.030	Latin America	0.060	0.037
Latin America	0.018	0.026	Australia	0.194	0.019
Eastern Europe	0.001	0.012	Eastern Europe	0.022	0.015
India	0.001	0.007	Other Asia	0.022	0.007
Australia	0.014	0.006	Korea	0.005	0.004
Thailand	0.001	0.006	Thailand	0.000	0.002
Other Asia	0.000	0.005	India	0.001	0.000
Rest of the world	0.007	0.041	Rest of the world	0.090	0.009
Top 5 (overall)	**0.994**	**0.988**	**Top 5 (overall)**	**0.979**	**0.992**
Top 4 (non-US)	**0.961**	**0.903**	**Top 4 (non-US)**	**0.717**	**0.906**

Source: U.N. Comtrade data set (online)

Mexico, the second most important trading partner for Canada in 1993, is now surpassed by China, where more than half of all Canadian (non-US) parts exports were heading in 2003. The large increase of exports to China, where the domestic automotive industry is expanding rapidly, seems particularly vulnerable. For example, Magna International, by far the most important Canadian parts exporter, is increasing its production capacity in Asia. These trends are already reflected in the huge relative fall-off of Australia and the rest of Asia as an export destination for Canadian parts. In contrast, the most rapidly increasing assembly regions of the world, except for China, i.e. Eastern Europe, Korea, Thailand, and India, are not yet important trading partners for Canada.

On the import side, on the other hand, the growth of the automobile industry in developing countries is already making a small impact. The share of Canada's part imports coming from the top 4 countries (excluding the U.S.) decreased from 96.1% to 90.3%. The countries with growing automobile industries figure prominently. China, Eastern Europe, India, and Thailand all post enormous increases, albeit from a low base.

Finally, we take a look at the 10 most important (6 digit HS) products in Canada's exports; Table 1.3 has the list. These account for 93.4% of Canadian parts exports in 2003. The share of the United States is again extremely high and for most products the U.S. import growth is rather high. The right-most columns indicate the export growth of Mexico and China in each of the ten parts. The export levels for each of these parts are relatively low in these countries, but the growth rates are extremely high. Importantly, they tend to be much higher than the growth rate of U.S. import demand.

For the Canadian industry it is extremely important to continuously find new products where it can establish a comparative advantage. For example, in 1998 U.S. imports of "fittings" and "electric lighting" were negligible and currently all Canadian exports of these two products go to the U.S. In 2003, these two products combined accounted for more than 18% of Canadian parts exports to the U.S.

Table 1.3: The 10 most important Canadian automotive component exports in 2003

| Top 10 parts (excluding n.e.c.) | Share in Canadian parts exports | Importance of USA | | Export growth of | |
		Share	Import growth (av. +32%)	Mexico (av. +53%)	China (av. +307%)
Brake system	0.252	0.96	32%	42%	207%
Bumpers	0.165	0.96	18%	158%	121%
Fittings	0.136	1.00		74%	-71%
Wheels	0.126	0.91	79%	157%	554%
Mufflers	0.069	0.99	79%	281%	1835%
Safety glass	0.048	0.87	23%	-6%	225%
Electric lighting	0.046	1.00		199%	264%
Shock absorbers	0.036	0.89	20%	1090%	3066%
Safety belts	0.033	0.73	-44%	14%	4848%
Seats	0.023	0.99	-99.7%	301%	2401%

Source: U.N. Comtrade data set (online)

1.2 Vertical organization of the industry

1.2.1 Current impact

Until the 1960s, the two major firms in the North American industry, GM and Ford, were highly vertically integrated. Chrysler outsourced a larger fraction of its component inputs. The establishment of Japanese-owned assembly plants on the continent, starting in 1982, reversed the trend and independently owned suppliers flourished. Managing a supply chain of several thousand firms proved to be exceedingly complicated and over the last 15 years the industry organized into a tiered supplier network. The final vehicle producers—OEMs—would out-source major components or subassemblies to Tier 1 suppliers, which in turn outsourced several of the components to Tier 2 suppliers, and so on. As a result, the OEMs dramatically re-

duced the number of suppliers they had to deal with directly, without giving up the benefits of specialization.

1.2.2 Future impact

In recent years this arrangement has started to change again. It is too early to know whether these trends will radically change the organization of the industry, but the following five issues have received a lot of attention recently:

- Suppliers are increasingly involved in the design and development of the parts they produce. One of the main benefits is to cut development time. While twenty years ago, a model had an average product life of 7 years, the main Japanese producers now introduce new versions of the majority of models in their lineup every 4 years. Hyundai is even trying to achieve the same feat every three years. To facilitate this rapid product turnover, R&D is pushed upstream. In 2003, Andrew Brown, Delphi's executive director of engineering, claimed his company was spending US$2 billion in R&D and engineering worldwide, almost 8% of sales: "Most innovations in safety, emissions, and entertainment come from Tier 1 suppliers." In a 2002 study prepared for Accenture by the Center of Automotive Research (CAR) in Michigan, the share of components in the total value generated in the U.S. automobile industry was estimated at 58.3% for 1990, against 24.5% of the value generated by the vehicle producers. This declined to 56% in 2000, but is expected to increase to 63-65% by 2010.
- Cost control by OEMs is increasingly focused on streamlining the supply chain. The process of outsourcing entire modules to Tier 1 suppliers and delegating responsibility for the design and subcontracting has probably gone furthest in interiors and seats. Lear, Johnson Controls, and Intier dominate that industry and handle the design of complete vehicle interiors. Recently, GM announced that it would take more control over its interiors and work directly with smaller suppliers. GM believes it can more effectively control costs and quality by bringing more work in-house. This initiative is just one facet of wide-ranging cost cutting programs in purchasing that all the major auto-

makers have engaged in. GM has just finished a three year program aiming to cut its component costs by 20%. Given that the company's purchasing bill runs at US$86 billion annually, savings can potentially be huge. An important new addition to the program is that in the next two years GM will require all its 250 largest suppliers to have offshore manufacturing capabilities. This is in addition to any price target.

- While assemblers are bringing some tasks back in-house, at the same time the role of Tier 1 suppliers is increasing in some vehicle programs. For niche vehicles or low-volume cars the entire assembly is sometimes turned over to an outside contractor. This practice allows OEMs to assemble vehicles locally without large capital investments or to increase production capacity when their own assembly plants cannot satisfy demand for an unexpectedly successful model. In addition, suppliers are sometimes in charge of building a convertible or stretched vehicle from an existing sedan or adding four-wheel drive. Magna Steyr is a prime example of such a "Tier 0.5 supplier" strategy, with an increasing focus on assembly. Currently it produces the Mercedes-Benz G-class, Jeep Grand Cherokee, and the Chrysler 300C in Europe for DaimlerChrysler and it is the sole assembler for the BMW X3. In the past it also developed and produced convertibles, four-wheel drive, and stretched vehicles for Saab, Volkswagen, Audi, and Mercedes-Benz and it is currently designing the new Stillo for Fiat. Karmann in Europe and ASC in North America are other firms with expertise in this area.

- An alternative to outsourcing the assembly entirely is to bring modulization to the assembly plant. An important trend, especially in Europe and Latin America, is the factory-within-a-factory cooperation between OEMs and suppliers. Within the Nissan assembly plant in Sunderland (U.K.), Karmann installs the folding hardtop roofs on the Micra. Starting in 2006, when the new compact minivan will be introduced, Magna Kansei will install its own cockpit modules and Calsonic Kansei the front-end modules, again operating within the Nissan plant. Similarly, Kuka Group will run the paint, body, and chassis operations of the new Daimler- Chrysler Toledo (OH) assembly

210

plant, which will open in 2006 and produce the Jeep Wrangler. Kuka and three other suppliers are investing US$300 million in the new plant. Other important projects are DaimlerChrysler's Smart plant in France, its Campo Largo plant in Brazil, Volkswagen plants in Resende, Brazil and Mlada Boleslav, Czech Republic and GM's Blue Macaw plant also in Brazil.

- Finally, the closer integration of OEMs and their suppliers increase the stakes when unexpected things happen, such as the current spike in raw material prices. This has led to a number of bankruptcies and court cases. With several large suppliers, most notably Delphi, Collins & Aikman Corp, Tower Automotive, and Federal-Mogul in Chapter 11 bankruptcy restructuring, the exposure of OEMs to problems at their suppliers is becoming apparent. Without Ford's assistance Visteon would also have had declare bankruptcy. Several of the companies are kept alive by credit from their clients which would suffer from the disruption of their supply chains. Several disputes center around the sharing of increased raw materials costs. While the most successful suppliers, such as Robert Bosch and Valeo, have been able to pass some of the increases in steel prices to their clients, a similar attempt by Lear has landed it in court. Its dispute with DaimlerChrysler affects 12 final assembly plants and is closely watched by the rest of the industry.

1.2.3   Canada's position

The share of intermediate inputs as a fraction of the value of vehicle production is higher in Canada than in the United States. Figure 1.7 uses industry data compiled by the OECD to track the evolution of the material cost/sales ratio for the motor vehicle industry in each of the three North American countries. The Mexican and, especially, the Canadian industries outsource more of their material purchases than the U.S. industry. In Canada, material purchases as a percentage of final sales in the motor vehicle industry even exceeded 80% in the early 1970s, while in the U.S. it peaked at 69%.

From 1971 to 2001, assembly plants in each country initially outsourced more tasks and inputs, but this reversed

211

towards the end of the century. In 2001, the U.S. industry is back where it started, purchasing 60% of its sales value from other industries. The Mexican ratio converged to the U.S., testament of its close integration in recent years. In Canada materials still take up a larger share of sales, which might be related to the important presence of DaimlerChrysler and Magna, two companies that have been instrumental in the push towards modulization of assembly.

Figure 1.7: Share of intermediate inputs in total sales
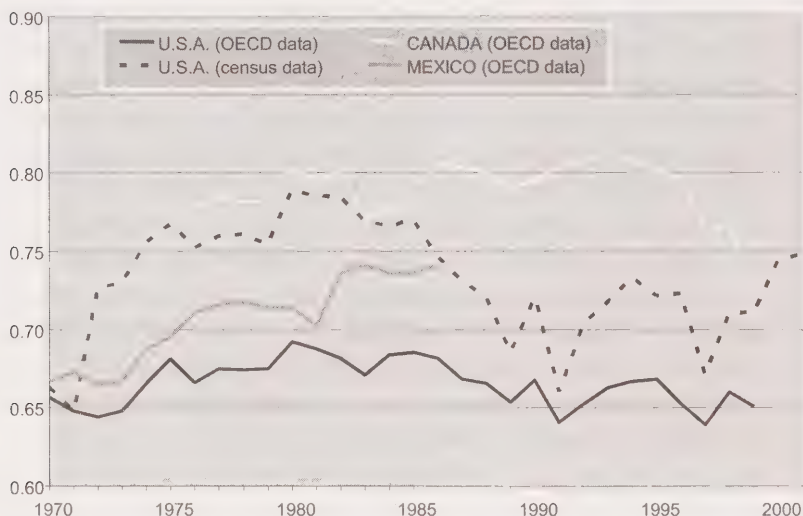
Source: STAN data set (OECD) and U.S. Bureau of the Census

Table 1.4 contains the most important Canadian firms in the automotive parts and components sector. The importance of Magna International is striking. It is more than 10 times larger than the second Canadian firm, Linamar and 18 times larger than the ABC Group, the Canadian number three. The Canadian share of North American light vehicle production in 2002 was almost 16%, but only 5% of the major component suppliers have Canadian headquarters. In 1999, Faurecia, Decoma, and F&P Manufacturing still were operating regional headquarters in Canada, but by 2002 their Canadian affiliates did not report as separate suppliers anymore.

212

Recently, wage pressures on the industry have increased. The large incentives offered by the OEMs to boost sales have been accompanied by increased cost cutting efforts. Outsourcing components to Asia, especially to China, is a first manifestation of this trend. The difficulty of Delphi and Visteon, formerly owned by GM and Ford, to maintain high wages is a second manifestation. The 2003 wage negotiations between the UAW and Delphi introduced dual wage profiles at the largest North American supplier, allowing the firm to pay newly hired workers less than insiders. Finally, given that suppliers tend to be smaller firms than OEMs, the wage gradient by firm size influences the relative competitiveness of different countries in attracting suppliers. Statistics in Table 1.5 indicate that, even though the average salary in the automotive industry was significantly lower in Canada than in the U.S., this is reversed for the smallest firms. U.S. firms that employ less than 20 employees paid an average salary of C$30,940 in 1995, while Canadian firms of similar size paid C$36,300.

Table 1.4: Canadian top suppliers (NA rank)

	1993 (top 100)		1999 (top 150)		2002 (top 150)	
	sales	NA rank	sales	NA rank	sales	NA rank
Magna Int.	2,450	7	5,760	6	7,650	5
Linamar Corp.			687	54	712	56
AG Simpson			407	91	245	129
Multimatic			356	107	342	110
ABC Group			323	110	423	94
Meridian Tech.			306	117	207	147
FAG Autom.					210	145
Fabricated St. P.	160	71				
Faurecia			586	64	HQ in France	33 (parent)
Decoma Int.			496	80	HQ in US	45 (parent 2000)
F&P Mfg			291	125	HQ in US	84 (parent)
Source: Automotive News						

213

Table 1.5: Wage gradient by employment category in the motor vehicle industry (1995)

Employees:	1-19	20-99	100-499	499+	Industry average
Canada	0.75	0.71	0.81	1.1	C$48,400
USA	0.55	0.64	0.68	1.06	C$56,250 (US$41,000)

Note: Average salary (including benefits) for firms in different size-categories (measured by employment) as a fraction of the average salary for the industry.

Source: OECD

Of course, comparisons like this are highly sensitive to the exchange rate of the moment. The enormous appreciation of the Canadian dollar since 2002 has further eroded the competitiveness of Canadian suppliers.

While the automobile industry spends enormous sums on R&D, to a large extent this bypasses the Canadian industry. The vast majority of innovation is done at company headquarters, which are only rarely in Canada. Table 1.6 lists the location of the headquarters of the top 150 (top 100 in 1993) suppliers in the North American automobile industry. While 9 of the largest firms were headquartered in Canada as recent as 1999, this declined to only 7 in 2002. Only two companies improved their rank. At the same time, Mexican and European firms increased their presence. In contrast, the importance of Michigan is striking. In 2002 it was home to 82 of the top 150 suppliers, a full 55%. Within Michigan, firms are concentrated in Detroit, Troy, and Auburn Hills, where GM, Ford, Chrysler, and Delphi have their headquarters. As the OEMs have disintegrated vertically, geographic proximity has become a substitute for ownership ties to smooth commercial interactions.

Table 1.6: Location of top supplier headquarters (to the NA industry)

	1993 top 100	1999 top 150 (100)	2002 top 150 (100)
Canadian	2	9 (5)	7 (3)
Mexican	0	2 (1)	3 (3)
European	0	2 (2)	4 (3)
Michigan	51	72 (51)	82 (52)
Detroit-Troy-Auburn Hills	24	32	33
Other Midwest[1]	24	36	30
Southern U.S.[2]	8	17	16

Notes:
[1] OH, IN, IL, and PA;
[2] AL, FL, KY, MO, MS, NC, SC, TN, and TX

Source: Automotive News

1.3 Flexible production

1.3.1 Current impact

In the automobile industry, flexibility is the new buzzword in manufacturing. Increased competition has led manufacturers to increase the number of products they offer for sale. Traditionally, each assembly plant produced a single model or a few similar ones. The explosion in models for sale made it prohibitively expensive to continue this practice. As a result plants are being forced to assemble several models on the same assembly line, which has important consequences for production and trade.

A first effect is that the exploitation of flexibility often goes together with diminished emphasis on realizing scale economies. While the average plant size has been decreasing gradually over the last 30 years, as the industry made the transition from mass to lean production, the recent decrease is more pronounced[4]. New plants announced in North America have often been in the 100,000 unit range, although subsequent capacity

---

[4] For detailed information on this, see Van Biesebroeck (2006), "Productivity Dynamics with Technology Choice: An Application to Automobile Assembly," *Review of Economic Studies*, 70(1), pp. 167-98.

215

additions have made the difference with existing plants smaller. China, where most of the recent capacity additions have taken place, has an average plant size of approximately 50,000 vehicles[5]. The minimum efficient scale of an automobile assembly plant seems to be falling.

It is hard to know whether it is the cause or the effect of the manufacturing flexibility, but the number of models for sale has increased dramatically in recent years. Table 1.7 illustrates this trend over the last 30 years. Models for sale in the United States increased from 133 in 1974 to 282 in 2004. The growth has been much more pronounced in light trucks than passenger cars, and in the former category it does not seem to have topped out. The trends for the number of models sold and produced in North American are by and large similar.

Table 1.7: Number of car and truck models sold and produced in North America (1974-2004)

	1974	1984	1994	2004
Models for sale in U.S.	133	195	238	282
Cars	96	140	164	167
Light trucks	37	55	74	115
Models for sale in NA	185	228	273	320
Models produced in NA	90	125	139	165

Source: Ward's Automotive Yearbooks and Ward's Infobank (2004)

One way to increase product variety is to sell mechanically similar cars under different nameplates. Models that share a platform can be made to differ mainly in appearance, standard features, and trim level, while it is straightforward to develop and assemble them together. All firms have mastered such a 'platform stretching' strategy, even though they do not all use it to the same extent. The number of platforms in production has increased notably less than the number of models[6].

---

[5] The substantial involvement of provincial governments in China, aimed towards attracting automotive investment to their province, makes it not unlikely that production in China is taking place below efficient scale.

[6] For details, see Van Biesebroeck (2006), "Complementarities in Automobile Production," *Journal of Applied Econometrics*, forthcoming.

216

A by-product of the increased variety is the emergence of new market segments and different sources of differentiation. Imported vehicles are no longer either small, reliable and cheap or high quality, luxurious sport sedans. Given the multidimensional product competition, the nationality of the owning firm is becoming an ever smaller factor to explain a vehicle's attractiveness. From the consumer's point of view, the difference between vehicles produced by domestic or foreign producers is becoming smaller. Their product lines overlap almost completely and there is no single characteristic on which domestic and foreign vehicles differ consistently.

1.3.2   Future impact

Smaller, nimbler plants could be operated at a higher rate of capacity utilization. Recent research of CAR in Michigan indicated that, in the latest economic downturn, capacity utilization in the industry hardly declined, even though profits were dragged down by lower prices. It is suggested that the break-even point in capacity utilization has increased substantially over time. Trade can be a contributing factor to make sure factories operate as close to full scale as possible. Previously, plants had to be dedicated to a single model and capacity utilization fluctuated with the popularity of that model. The ability to produce a wide variety of models in a single plant allows firms to tailor production more closely to demand. Especially for foreign producers, flexibility allows firms to rely less on imports and produce more domestically, operating their North American plants closer to full capacity.

The average plant size in countries with more recently built assembly plants is certainly lower than in the United States or Canada. Even in Mexico, average capacity is 140,000 vehicles relative to 200,000 further north. In China especially, many smaller plants are being built. To some extent this merely represents cautious entry in an uncertain market or by new firms, but also more established plants by Western multinationals tend to be smaller. It is not implausible that Canadian plants will also become smaller in the future.

217

1.3.3 Canada's position

As mentioned, in the long run it is not impossible that greater flexibility will lead to lower imports even without new capacity additions. For example, Honda claims all its assembly plants can produce its entire model range with a relatively low productivity penalty. When the firm initially established a manufacturing presence in the United States it was only natural to produce first only its best selling vehicles, the Civic and Accord. Now that it operates several plants across the continent, flexibility will allow the firm to shift production to the vehicles most in demand and avoid having idle capacity in North America, while importing different models from overseas. However, for this to be a reality, the entire supply chain has to become equally nimble. Different vehicles require different components and the suppliers have not yet matched the OEMs' flexibility.

Even though Japanese plants in the United States were the first to be flexible, the technology is now spreading through the industry. In Canada, Honda claims to be able to assemble almost its entire line-up in each plant. Its Alliston plant in Ontario has produced a wide range of vehicles in the last decade. The Ford plant in Oakville is currently undergoing a $1 billion investment project to make it one of its most flexible facilities. The DaimlerChrysler plant in Windsor assembles three models derived from two different platforms, which is the ultimate in flexibility. Finally, also Toyota has manufactured a wide range of models in Canada, including the first Lexus being produced outside of Japan.

1.4   Stock of vehicles

1.4.1   Current impact

The North American light vehicle market recently has had a string of record sales years. At the same time the average expected lifetime of a vehicle in Canada has risen from 7.7 years or 154,000 km in 1970 to 11.6 years or 227,000 km today. As a result, new sales outnumber the number of vehicles that are scrapped, increasing the number of vehicles on the road. Figure

1.8A shows for the United States sales and registrations from 1972 to 2002, both normalized at 100 in the first year. Clearly, the combination of increased durability with record sales has increased the number of registrations ever higher. In 2002, almost 236 million vehicles were registered in the United States.

Figure 1.8A: Vehicle sales and registrations in the United States (1972 = 100)
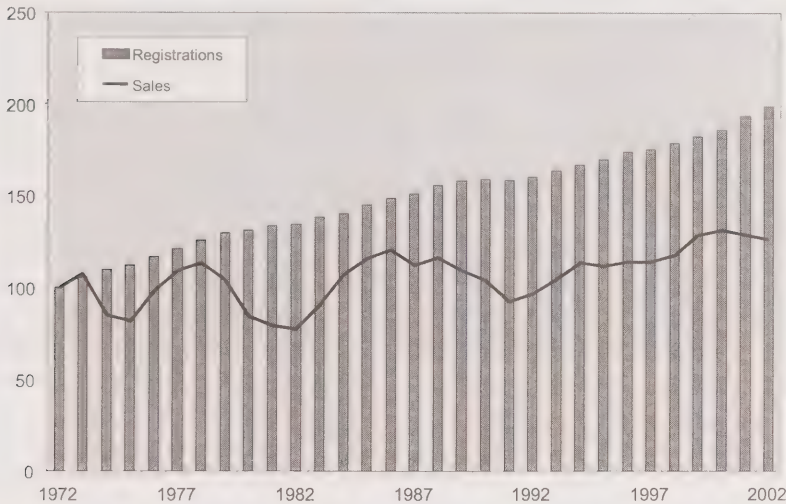
Source: Ward's Automotive Yearbook (various years)

To put the increasing stock of vehicles in perspective, Figure 1.8B plots registrations as a fraction of the population of driving age. For the United States this increased from 55% in 1960 to 102% in 2004. The growth in this ratio is projected to decelerate and only reach 103% by 2010, but that will only happen if sales of new vehicles drop far below current levels. The ratio is lower in Canada, but one cannot automatically infer that the potential demand is larger. Canadians are not as rich as Americans, on average, and more likely to live in cities, which lowers demand for vehicles. By 1990, there were 0.69 vehicles per person of driving age and this has remained virtually unchanged in the last 15 years, only reaching 0.70 in 2004. Growth opportunities in Mexico are much larger. The current vehicle penetration rate is much

lower and still increasing, although only very slowly, because population growth is relatively high.

Figure 1.8B: Total vehicle registrations per driving age population
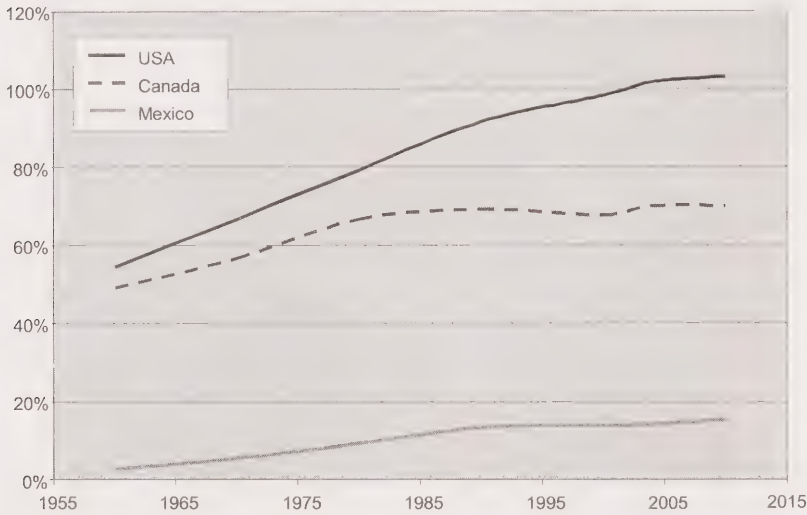
On the demand side quality is becoming a less important factor as well. This is not really surprising as there are decreasing marginal returns to everything. The large quality improvements by GM over the 1990s have not provided the anticipated sales boom. The stellar quality record of GM's Oshawa 2 plant did not prevent an announcement of its prospective closure. The quality record of Buick hardly translates in higher sales. In the 2004 Vehicle Dependability Study by J.D. Power which looks at longer term (3 year out) defects, Buick was the second most reliable brand in North America, only topped by Lexus. GM brands with average number of defects below the industry average include Buick, Cadillac, Chevrolet, GMC, Saab, and Saturn. Only Pontiac and the discontinued Oldsmobile perform more poorly than average, but this has not prevented GM's market share from slumping continuously. Similarly, Hyundai passed Toyota in the initial quality survey (after 3 months of ownership), but it still sells its cars at a discount relative to its Japanese competitors.

1.4.2 Future impact

Automobiles are durable goods and sales predictions are obviously affected by the stock of vehicles in the economy. The preceding analysis points to weaker sales in years ahead. At a micro level, the industry got a taste of what is to come when sales of the Big Three collapsed in the fall of 2005 after a summer where "employee discount plans" spectacularly increased sales. At an economy-wide level, the bumper sales years of the last half decade will most likely translate into lower sales years ahead. For example, Automotive News predicted 2006 sales to be 4% below the 2005 level for the U.S. at 16.5 million units. Ward's predictions were similar and they expect further declines in 2006 if GM and Ford hold firm on their commitment not to boost fleet sales[7]. Moreover, the mix of vehicles is also shifting towards more economical and smaller cars. For Canada, EDC Economics predicts exports of vehicles to decline by 3% in 2006 and 5% more in 2007, reflecting softening demand in the U.S.

At the same time the fleet is aging. As people owning a second hand vehicle are less likely to trade it in for a new car, future demand for new vehicles might fall off even more rapidly than the registration statistics suggest. Currently, the group of cars 1–5 years old is larger than the group of 10+ year old vehicles, and this is expected to remain true for another 3 years. Later, the group of very old vehicles will become the largest. As vehicle durability is maintained or even rises in the future, the owners of very old vehicles will have very little incentive to replace their vehicles. The solid lines in Figure 1.9 indicate that the resale value of four year old passenger cars and light trucks has decreased almost continuously over time.

The economic outlook of the government's Consensus Revenue Estimating Conference by the Administration, House Fiscal Agency, and Senate Fiscal Agency as agreed to at the January 12, 2006 meetings was slightly more optimistic forecasting 16.7 light vehicle sales in 2006 and 16.8 in 2007.

Figure 1.9: Predicted and actual resale values of 48 month old vehicles (Canada)
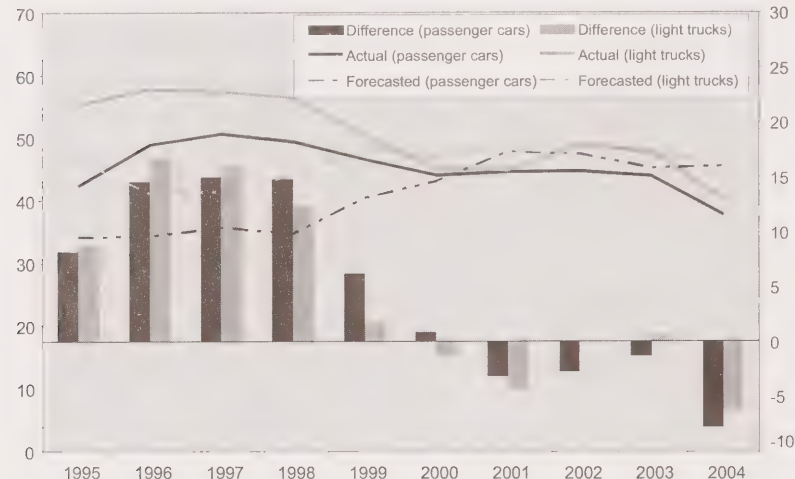
Source: DesRosiers "Industry Overview 2004" (2004)

In the past, firms promoted leasing to entice customers to trade in vehicles more quickly. In Canada, leasing peaked in 1997 at 46.8% of all new car registrations. In the following years, rising income levels made car ownership more affordable and leasing rates declined to only 28.1% in 2003. This suggests there is scope for growth by pushing leasing over buying. However, increased durability of vehicles has pushed up predicted resale price of off-lease vehicles, the dotted lines in Figure 1.9, making leasing very advantageous to consumers.

By 2004, lessors had predicted much higher resale value for their fleets than actually realized, resulting in negative off-lease values for their customers (the bars in Figure 1.9). This made it exceedingly disadvantageous for the lessees to take possession of their leased vehicles at the end of the contract, depressing recent resale values even further. Future lease contracts are expected to become more expensive as lessors take into account lower projected resale values. Less leasing could lead to slower vehicle turnover and lower sales of new vehicles.

Exporting excess supply of second hand vehicles to less developed economies is a viable alternative that E.U. countries are actively taking advantage off. Canadian exports of used

vehicles rose from around 15,000 in 1994-96 to more than 200,000 in 2001-02. The recent increase in the exchange rate has choked this trade. In the future, trading second hand vehicles with Mexico or other countries in Latin America could be an option that would benefit the Canadian industry. To stimulate the local automotive industry several countries, notably Brazil, have made trade in used vehicles very difficult. It is common practice for countries, even those with no domestic automobile industry, to charge higher import tariffs on second hand than new vehicles (often for emissions or safety reasons). The Free Trade Area of the Americas could prove very beneficial in this regard. Facilitating exports of second hand vehicles to the south would benefit the domestic industry.

1.4.3 Canada's position

What makes the previously described situation precarious for Canada is that the North American industry is plagued by overcapacity. For the industry, total excess capacity is estimated at approximately 0.5 to 1 million units, but this combines larger excess capacity at some firms, most notably GM and Ford, and a projected capacity shortage at other firms, notably Honda and Toyota. Over the last several years, the Big Three American firms have taken capacity from the market, while transplants are building new plants and this process is likely to continue. Canada has benefited from this as Toyota, Honda, and Suzuki (in a joint venture with GM) now operate plants in Ontario. The recently announced closure of the GM plant in Oshawa and the elimination of one shift in another plant will be partly compensated for by the new plant Toyota will build in Woodstock, Ontario, close by its current Cambridge complex.

The reductions in capacity far outstrip the additions. Including the GM announcement, 6 assembly plants will have been closed in Canada between 1993 and 2007, while only Honda and Toyota (in two locations) have substantially increased production capacity. Ford is expected to announce the closure of at least four assembly plants in North America early in 2006. While the large investment in a flexible production system for the

223

Oakville plant bodes well for its future, the future of the St. Thomas plant is more uncertain.

Over the longer term, the industry is only viable if production capacity matches demand. If future sales in North America will be lower than today, more closures will be inevitable.

1.5   New technologies

1.5.1   Current impact

The primary new technology in automobile production is the flexible plant, discussed earlier. Important evolutions in vehicle technology are taking place in powertrains. In Europe, fuel efficient diesel engines are outselling gasoline cars. Direct injection has vastly improved mileage and lowered emissions. In addition, diesel engines tend to last at least 25% longer than gasoline engines. The catalysts in the cleanest diesel powered vehicles require sulfur-free fuel, which will only be available in North America in 2006.

In North America, the preferred way to achieve similar fuel efficiency is through hybrids. A battery pack is added to the vehicle, which is charged by a smaller combustion engine and by power-recycling technologies when the vehicle brakes. In stop-and-go traffic an electric engine provides (additional) acceleration power, while the gasoline powered combustion engine can function at optimal operating speed. On the highway, gas mileage in the two most popular hybrids, Toyota Prius and Honda Civic (4.2 and 4.3 l/100km), is comparable to the Volkswagen diesels in the Golf or Beetle (4.6 l/100km); in city driving the hybrids deliver superior mileage.

The second important trend in vehicles is the growing importance of electronics. This was clearly illustrated in 2004 when Robert Bosch became the largest component supplier in the world, and Siemens VDO was the fastest growing of the top suppliers. Both firms specialize in electronics. Visteon and Delphi, the two largest North American suppliers, are rapidly increasing their electronics division, which for Delphi is already responsible for more than 20% of revenue. Not only are the electronics-

intensive firms growing most rapidly, they also spend a lot on R&D. Siemens VDO, Hella, and the electronics division of Visteon report spending 9% of sales on R&D; Bosch is not far behind at 7.1%. This is well above the industry average of 4%.

1.5.2 Future impact

The future of drivetrain technology is likely to be the fuel cell, as the "hydro economy" develops. Each major automobile manufacturer is involved in developing fuel cell vehicles, which is certain to represent a much more dramatic shift for the industry. The outlook and especially the timing are highly uncertain, but Canadian industry is very active in this field. Ballard Power, headquartered in Vancouver, is considered to be one of the world leaders in fuel cell technology. It is already a supplier to the automobile industry and thus well placed for the future. Other Canadian companies active in the development of fuel cell technology for vehicles are Astris Energy, Cellex Power, and Zongshen Pem Powersystems.

While fuel cells are important for the long-term future, electronics will matter greatly in the years to come. For North American suppliers, the top three concerns are to (1) broaden their client base to include transplants, (2) get compensated for raw material price increases, and (3) expand in electronics. The first and third items are identified as the most import growth opportunities for domestic suppliers. It is estimated that the electronics content in the average vehicle will increase from US$2,250 in 2000 to US$3,850 by 2010. In addition, OEMs expect that by 2010 50% of all R&D—a large fraction of this in electronics—will be carried out by suppliers.

Finally, the importance of the Internet is also felt in this industry. On the consumer side, in Canada as in the United States, new vehicles have to be sold by dealerships. As a result, online purchasing has never taken off, even though the second hand car market has taken advantage of the Internet to organize classified ads, but also for transactions. A result of the wealth of information accessible on the Internet is the increased bargaining power of customers, at the dealers' expense. Profit margins in dealerships have

declined noticeably in recent years. The "employee pricing" schemes that the Big Three ran in the U.S. and Canada over the summer of 2005 could have a lasting effect as consumers were particularly attracted to the no-haggling buying process.

On the B2B side of the market, the demise of the cooperative online auction website Covisint has left each company organizing much of its own purchasing again. For Canadian suppliers to OEMs as well as for the Canadian aftermarket it is especially important to follow developments in the United States. E-commerce applications are subject to network effects and getting locked into an incompatible standard can be very costly. At the same time, timely and accurate communications can provide large productivity gains and Canadian firms do not want to come late to this technology.

1.5.3 Canada's position

No Canadian assembly plant produces hybrids, and this is likely to remain so until Ford brings the hybrid versions of the Edge to Oakville, which is currently projected to happen only in 2010. While most Civics sold in North America are produced in Alliston, Ontario, hybrids are imported from Suzuka, Japan. The Honda Accord is also only produced with gasoline engines in Marysville, OH, while hybrids are imported from Sayama, Japan. Similarly, the Lexus RX 330 is produced in Cambridge, Ontario and Kyushu, Japan, but the Japanese plant is the only one that produces the hybrid version (400h). Even GM has chosen to launch production of its Chevrolet Silverado hybrid pickup truck in Fort Wayne, IN, even though the Oshawa plant is the lead plant for the vehicle. Ford produces its Escape hybrid alongside the regular Escapes in Kansas City, MO and Avon Lake, OH[8].

The current popularity of hybrids, also in Canada, is very strong. In 2005, the Honda Civic hybrid was chosen as the family sedan of the year by Consumer's Report while the Toyota Prius took top honours as the Car of the Year. Waiting lists

---

[8] It should be noted that GM does produce E85 vehicles in Canada, such as the Monte Carlo, Impala and the Silverado.

for the Prius are still running over half a year. North American sales in 2003, for all hybrids combined, were only 40,000, but are expected to reach 177,000 in 2005. A total of 28 models—18 trucks and 10 cars—are expected to offer hybrid powertrain options in 2008. Thus far, all of this is bypassing the Canadian industry.

It is unlikely that the new diesel technology will ever be as popular as the hybrids, but the Canadian industry is again not very involved. Two of North America's largest engine plants are in Canada. Ford's Windsor plant has a capacity of approximately 600,000 engines, mostly V6's, but its future capacity utilization will depend on Ford's future restructuring plans. GM's St. Catherines plant used to be even larger, but high fuel prices have put the demand for V8 engines in doubt. Cylinder deactivation technology, such as that in the Impala, allows for substantial fuel savings and have proven to be popular. No diesels or hybrids are made in Canada.

Of the top Canadian suppliers, Magna has a sizeable electronics division, and also the ABC Group, Canada's third largest supplier, is heavily involved in electronics. All other large Canadian suppliers, Linamar, Multimatic, A.G.S. Automotive Systems, Meridian Technologies, and FAG Automotive, tend to have their comparative advantage in mechanics. Advanced technologies are equally important here, but the value added share of the vehicle is clearly shifting towards electronics.

Finally, we list the trading partners and products that Canada is running a trade surplus with in automotive components. Figure 1.10 illustrates that Canada is running a trade deficit with all its primary trading partners in components. The deficit is especially large with Japan, where Canada is importing 20 times as much as it is exporting. With several of the fast growing automobile producing countries, such as Brazil, South Korea, India and Thailand, Canada is also running a deficit. China is the one positive note, but the rapid expansion of the Chinese automotive industry combined with a deepening of its domestic supply chain puts much of those Canadian exports in doubt for the near future.

227

Figure 10: Net exports as a percentage of total exports

LA ● ROW ● Austra● China

● USA

● Thailand

● Brazil

● Japan

(export-import)/export

trade partners (ordered by average trade volumes)
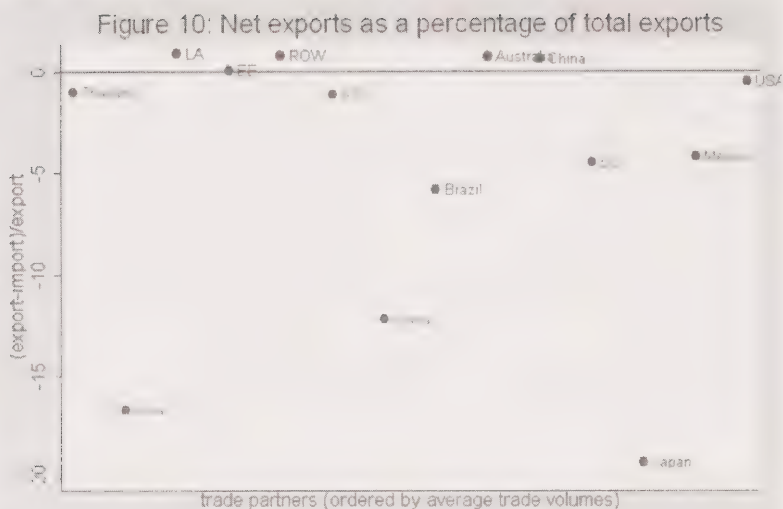
Source: Own calculations based on U.N. Comtrade data set (online)

Figure 1.11 illustrates the size of Canadian trade deficits per component, where components are ranked according to the importance of the E.U. in U.S. and Japanese imports. This ranking is intended to capture the extent to which a component can be considered high tech. The vertical axis represents the trade surplus or deficit: (exports-imports)/exports. The good news is that 90% of Canadian exports are components for which the E.U.'s importance as a source of imports into the other advanced economies is higher than the E.U.'s median importance. The largest circles, which represent the size of Canada's exports, are to the right. This indicates that Canada is specializing in goods in which Europe is a successful exporter to the U.S. and Japan, presumably "high tech" goods. It is also clear that Canada's

exports are highly concentrated. There are only two goods for which Canada is running a sizeable trade surplus, the two largest circles above the zero line. These are non-electrically powered work trucks (the left-most dot) and bumpers, the largest positive observations more or less in the centre. In all electronics, Canada is running a trade deficit.

228

Figure 11: Net exports by product as a percentage of total exports
(Size of the circle proportional to total exports)

Y-axis: (export-import)/export, truncated at -20

X-axis: Automotive components (Ordered by importance of EU in US/JAP imports)
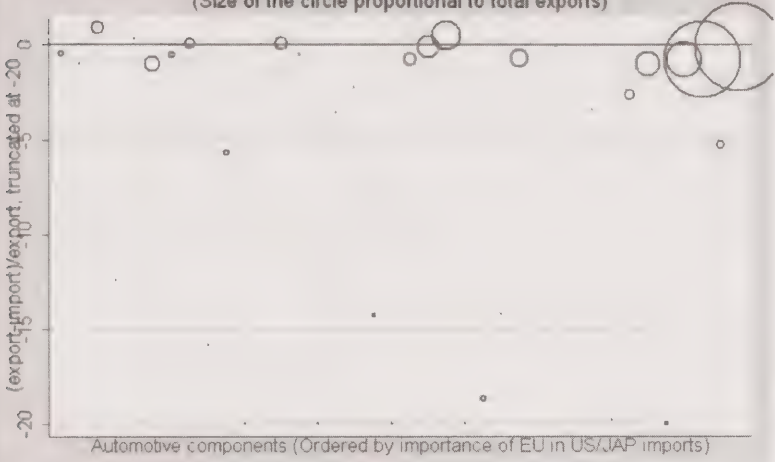
Source: Own calculations based on U.N. Comtrade data set (online)

229

2. Market analysis: automobiles and light trucks (with Frank Verboven)[9]

This section estimates the potential impact of eliminating MFN tariffs on new vehicles on the production, employment, consumption, and trade of new vehicles in Canada (both short and long-run effect) using cost-benefit and regression analyses (such as estimating the price and substitution elasticity, taking into account quality and reliability differences) with disaggregation by vehicle type to the extent possible. The estimation will be undertaken under two scenarios:

a) Unilateral elimination of Canadian tariffs on new vehicles;

b) Elimination of tariffs on new vehicles in the following five FTA contexts: Canada-South Korea, Canada-E.U., Canada-Japan, Canada-China, and Canada-Mercosur (each of the five separately).

The expected effect of eliminating the 6.1% tariff on final vehicles can be broken down into the following components:

Effect = Benefit – Cost

=	**Lower price for consumers**	**(1)**
	+ higher sales of vehicles	**(2)**
	+ tariff concessions by trade partners	**(3)**
	– lost tariff revenue	**(4)**
	– lost FDI	**(5)**
	– lower domestic production	**(6)**

Items number 1, 2, 4, and 6 will be addressed in this section[10].

Item number 5 is the subject of the next section.

---

[9] All analysis in Section 2 is joint with Professor Frank Verboven from the Catholic University of Louvain in Belgium.

[10] Note that this analysis is limited to final assembly. If Canadian parts are disproportionately oriented towards Canadian vehicle production, the lost production domestically will have a multiplier effect on the parts sector. Given that we do not have any data on this exposure, we merely note this point, but do not come up with an estimate of the effect.

Properly discussing item number 3 would go beyond the automotive industry. We will include some comments about the possible export effects of trade concessions of trade partners on final vehicles in the current section, but reciprocity in the automobile industry is only one of several possibilities.

Throughout, we will have to compare dollar values that represent gains and losses for different groups to obtain an aggregate effect for Canada. When an effect has a non-obvious distributional effect, it will be noted.

The way we will obtain estimates for the quantities outlined above is by estimating a discrete choice model of vehicle choice in the Canadian automobile market. This follows in a recent tradition of using oligopolistic models of competition in differentiated products to study the actual market equilibrium and to conduct counterfactual analysis. The crucial objective is to get an estimate of the primitives of the model, most crucially the demand parameters, but potentially also the parameters that govern the marginal cost function. With estimates for those functions in hand, one can conduct counterfactual simulations how the market equilibrium is expected to change if, for example, a trade policy is changed. The main benefit of such an approach is that we allow all market participants, even those only indirectly affected by the policy change, to update their strategies and we calculate a new Nash equilibrium for the industry. This way one obtains a consistent estimate of the trade policy effect only keeping the primitives constant, not the observed strategies. It leads to an analysis that is robust to the Lucas-critique, which has plagued earlier counterfactual analyses.

We will proceed in three steps. First, in Section 2.1, we formulate and estimate a discrete choice model of vehicle demand. Given the time constraints for this project we estimate a nested logit model using aggregate market shares and model characteristics, including price, at the vehicle-level. In a more elaborate analysis one could allow random coefficients on some of the characteristics, especially price, and add assumptions on the shape of the marginal cost function to estimate a supply equation jointly with demand. In Section 2.2, we outline our estimation strategy.

231

Second, in Section 2.3, we use the demand model to calculate a number of quantities that are generally unobserved, but which will influence the effect of any policy change. In particular, we will calculate (i) own and cross-price elasticities for each model with respect to all other models in the market; (ii) unobserved vehicle quality, from the point of view of the consumer; (iii) the marginal costs for each vehicle that are consistent with the estimated price elasticities of demand and the observed prices. We do not estimate the supply side of the market directly, as it is not necessary to identify the demand parameters. It could result in more precise estimates and would allow one to impose the condition that firms always set prices on the elastic portion of demand, as theory implies. We will test how frequently this last condition is violated if it is not imposed. To calculate the elasticities and marginal costs we will assume that firms are playing a Bertrand price-setting game in differentiated products. We will take explicitly into account that firms that produce multiple models will internalize the effects of a price change of one model on the sales of all their other models.

Third, using the estimated demand parameters, price-elasticities, and marginal costs we conduct counterfactual simulations of market equilibrium—in Section 2.4. In particular, we look at the impact of elimination of the 6.1% import tariff on non-NAFTA vehicles. This will take the form of a reduction in the marginal costs for the affected importers by 5.75% (as the calculated marginal cost includes the current tariff rate). Different scenarios for the extent of trade liberalization will change the models which are affected[11]. We calculate a number of summary statistics in each scenario to illustrate the impact on prices, mark-up, sales, production, profits, consumer surplus,

---

[11] Note that a few models produced in the U.S. do not meet NAFTA content requirements, e.g. the BMW X5 and the Mercedes-Benz M-class SUVs. We will not consider trade liberalization that eliminates this content requirement because it would severely complicate the analysis. While this is strictly speaking not consistent with full trade liberalization, one of the scenarios considered, these models are sold in sufficiently small quantities that we are confident it has only a marginal impact on the results.

232

tariff revenue, and the differential impact on domestic producers and importers[12].

2.1 Specifying a model of demand

The automobile industry has proved to be a popular proving ground for discrete choice models that estimate demand for differentiated products. The state-of-the-art in estimating aggregate demand is the random coefficients model discussed in Berry (1994) and first taken to the data (U.S. automobile purchases) in Berry, Levinsohn, and Pakes (1995). Micro-level data, as in Goldberg (1995) or Berry, Levinsohn, and Pakes (2004), can be used to obtain more precise parameters. An intermediate solution, in Petrin (2002), adds micro-moments to the aggregate estimation. Several studies have used these models to evaluate trade policies. Important recent studies that use aggregate data include Irwin and Pavcnik (2004) for airlines and Fershtman and Gandal (1998), Berry, Levinsohn, and Pakes (1999), Brambilla (2005), Brenkers and Verboven (2006) for automobiles. Section 2.5 contains a (non-exhaustive) list of papers that use discrete choice models to estimate the demand for automobiles. No estimates for Canada are currently available.

We will use a nested logit model; see Anderson and De Palma (1992) and Verboven (1996a) for details and Berry (1994) for a comparison with the general framework. This model can be interpreted as a restricted random coefficients model, see Cardell (1998), where consumers share the valuation on all the observable characteristics, except on a set of nesting dummies that segment the market.

Consider the Canadian automobile market where I consumers are considering to purchase a car or light truck. They can choose between J available models, one of which is the outside good, i.e. purchasing a second hand vehicle or postponing the

---

[12] Note that we have explicitly chosen not to calculate employment effects. As the results will make clear, quantity changes are relatively small. It would be entirely arbitrary to map these small quantity changes in employment changes because production is organized in large scale plants and indivisibilities matter.

purchase. The utility of the outside good purchase will be normalized to zero[13]. A consumer i's conditional indirect utility function from purchasing product $j = 1...J$ that belongs to nest/segment g is given by:

$$u_{ij} = \underbrace{\sum_{k=1}^{K} x_{jk}\beta_k + \xi_j - \alpha p_j}_{\delta_j} + \xi_{ig} + (1-\sigma)\varepsilon_{ij}$$

$$= \delta_j + \sum_g d_{jg}\xi_{ig} + (1-\sigma)\varepsilon_{ij}$$

Utility thus consists of a component that is common to all consumers (δ_j) which groups together the first three terms, a random taste of consumer i for vehicles in segment g (which can be positive or negative), and an individual-model specific random utility draw (ε_{ij}). The common part (δ_j) depends on K observable characteristics that each consumer values identically (fuel-efficiency, horsepower, size, etc.), a model-specific unobservable characteristic (combining the effect of style, advertising, etc.), and price—the only endogenous characteristic (which has a negative coefficient attached to it). The benefit of such a modeling strategy versus specifying a traditional demand system at the product level is that with only a few para-meters we are able to generate cross-price derivates between all models that are very general. Note that in 2005 a total of 238 different models were sold in the Canadian market. Specifying the demand directly would require an extraordinary amount of parameters to allow for flexible substitution patterns.

We assume that the distribution of the random utility term (ε_{ij}) follows the extreme value distribution, such that we can derive market shares in analytical form; for more details on the nested logit model see Anderson and De Palma (1992) and

---

[13] Note that to define market shares we have to define the potential market of consumers. With only two years of data, this decision is entirely inconsequential; it merely scales the market shares. We choose the number of Canadian households as our measure of I, which gives an market share for the outside product of almost 80%.

234

Verboven (1996a). We further assume the market can be partitioned into G exclusive and exhaustive segments. Each segment contains J_g models and $\sum_g J_g = J$. Each consumer will choose one model to maximize her utility.

The nested logit distributional assumptions on the random utility term yield the following choice probability for individual i for product j that belongs to segment h as a function of the entire $Jx1$ price vector:

$$s_{ij}(p) = \frac{\exp((\delta_j - \alpha p_j)/(1-\sigma))}{\sum_{l=1}^{J_h} \exp((\delta_l - \alpha p_l)/(1-\sigma))} \cdot \frac{\exp(I_h/(1-\sigma))}{\sum_{g=1}^{G} \exp(I_g/(1-\sigma))}$$

where

$$I_g = (1-\sigma)\ln\sum_{l=1}^{J_g} \exp((\delta_l - \alpha p_l)/(1-\sigma)$$

is called the 'inclusive value' for segment $g = 1...G$. The predicted aggregate market share for model j is obtained by averaging the choice probabilities over all individuals, which in our (simple) case is simply $N.s_{ij}$ because our choice probabilities are not individual-specific[14].

The nested logit model will result in higher elasticities of substitution between models in the same segment than across segments, which is a major improvement over the simple logit model. An unattractive feature is that the own-price elasticity of substitution for each model will be increasing in price. This will be discussed at length in the next section.

The model can be generalized in a variety of ways. Two approaches to add flexibility to the estimated own-price elasticities is to let the parameter that governs the degree of substitution within

---

[14] In the full random coefficients model, see Berry (1998), the market shares cannot be derived analytically because the choice probabilities vary by consumer. As a result, a simulation estimator has to be used and the unobservable quality term has to be calculated using an embedded contraction mapping. Both of these complications severely increase the computational burden.

nests (σ) vary by segment. If demand elasticity is higher for cheap small cars than for expensive luxury cars, it would show up as a higher σ parameter in the small car segment; see Brenkers and Verboven (2006) for an illustration on the European car market. In order to estimate this model we would require more data than we currently have. The severe time constraints on this project necessitated us to estimate a relatively simple model[15].

An alternative would be to introduce (more) random parameters to the model, which would allow different individuals to value the characteristics differently. The most direct way to obtain more realistic demand elasticities would be to let the coefficient on price to vary by income level. By simulating a sample of consumers with income levels drawn from the national income distribution, we can calculate the choice probabilities at a more disaggregate level. However, working with individual-specific choice probabilities would greatly increase the computational burden on the estimation because closed form solutions would not exist anymore and a fixed point iteration would be required to uncover the unobserved model characteristics. Berry, Levinsohn, and Pakes (1995) outline the approach.

2.2   Estimating the demand model

We estimate the nested logit model introduced in the previous section using seven nests: small cars, mid-size and large cars, luxury cars, compact and mid-size SUVs, large and luxury SUVs, minivans, and pickup trucks. We collected data on each model for sale in the Canadian market in the 2004 and 2005 model years[16]. Dropping all models that sell less than 200 units per year gives us a sample of 442 observations, 218 in 2004 and 224 in 2005. We have renamed some 2004 models because the replacement models were introduced under a different name, even though they are

---

[15] With only two years of data available, the substitution parameters σ for some of the nests were estimated to be (insignificantly) larger than unity, which violates the theory. Therefore, we forced them to be the same across nests.

[16] The model year runs from September 1 to August 31. This will avoid including observations in the sample where a vehicle is only sold for part of the calendar year.

clearly replacing an existing car in their segment. This affects only the estimates that use random (or fixed) effects.

Table 2.1 Summary statistics for Canadian (domestic) market

	Average	Standard Deviation	Minimum	Maximum
Price ($)	37,480	20,403	12,995	131,300
Sales (units)	7,030	10,919	203	70,853
Model characteristics:				
Hp/weight	0.551	0.135	0.249	1.262
Size (l x w x h)	0.882	0.220	0.357	1.452
Miles/$	2.322	0.819	0.952	7.048
Automatic	0.554	0.498	0	1
Foreign brand	0.567	0.497	0	1
Production location (for vehicles sold in Canada):				
Canada	10.3%	30.4%		
U.S. & Mexico	48.2%	50.1%		
E.U.	17.0%	37.6%		
Japan	15.6%	36.4%		
South Korea	8.9%	28.6%		
Segment:				
small car	14.3%	35.1%		
middle car	18.8%	39.1%		
upper car (large & luxury)	18.3%	38.8%		
lower SUV	17.9%	38.4%		
upper SUV (large & luxury)	13.8%	34.6%		
minivan	8.9%	28.6%		
pickup	8.0%	27.2%		

As explanatory variables, we follow most closely the papers by Berry *et al.* (1995) and Petrin (2002). The following variables are included: power is captured by horsepower per weight, size by length x width x height, and fuel efficiency by miles per dollar. We include a dummy variable that indicates whether an automatic transmission is part of the standard equipment as a measure of luxury; and a dummy whether the nameplate has traditionally been owned by a domestic producer. Note, for

example, that this latter variable is zero for the North-American produced Honda Civic—as Honda still tends to be perceived as a foreign car company. Similarly, all Volvos are foreign even though they are now owned by Ford and all Chevrolets are labelled domestic, even though some are manufactured by GM Daewoo in South Korea. Summary statistics are in Table 2.1.

Table 2.2 Demand coefficient estimates

	Dependent variable: logarithm of market share (relative to outside good)			
	OLS	Nested logit	Nested logit with IV	Nested logit with IV and RE
	(1)	(2)	(3)	(4)
Price	-0.037	-0.025	-0.052	-0.051
	(.003)***	(.001)***	(.003)***	(.004)***
Hp/weight	0.178	-0.198	1.648	0.823
	(0.511)	(0.220)	(.347)***	(.213)***
Miles/$	0.273	0.216	0.089	-0.003
	(.107)**	(.046)***	(0.065)	(0.061)
Size	1.196	0.075	0.454	-0.005
	(.443)***	-0.194	(.270)*	(0.222)
Automatic	-0.446	-0.026	0.203	0.177
	(.132)***	(0.058)	(.094)**	(.074)**
Domestic	-0.373	-0.133	0.050	0.020
	(.117)***	(.051)***	(0.077)	(0.092)
Nesting variable		0.859	0.693	0.698
		(.020)***	(.064)***	(.090)***
Year	0.000	0.054	0.037	0.020
	(0.102)	(.044)	(0.060)	(0.019)
Constant	-7.582	-3.580	-4.526	-3.474
	(.707)***	(.318)***	(.504)***	(.523)***
Observations	441	441	441	441
Adj. R^2	0.412	0.891	0.802	0.822

Notes: * Significant at the 10% level, ** at 5%, *** at 1%.

Parameter estimates for the demand system, using several estimation methodologies, are in Table 2.2. The simple least squares estimates, results in column (1), indicate that people positively value more engine power, greater fuel efficiency and a larger size. Contrary to expectation, a standard automatic is valued negatively and the willingness to pay for domestic cars is significantly lower than for foreign cars. Not surprisingly, consumers prefer paying a lower price, although the point estimate on the price variable is relatively low. Such a low coefficient estimate, -0.037, would indicate pricing on the inelastic portion of demand for a number of models, which is inconsistent with profit maximizing behaviour.

Coefficient estimates for the nested logit model, in column (2), are largely similar. Only the willingness to pay for horsepower turns negative as well, although insignificant. The parameter on the nesting variable is estimated to be positive and below one, in line with economic theory. The implication is that the cross-elasticity of price for models in the same nest is significantly higher than between models in different nests. This captures that consumers are more likely to substitute between models in the same nest; i.e., the segment classification that the industry usually employs makes economic sense. More worrying is the even lower estimated price coefficient than in the first column.

While the low estimates for the price coefficient, and the low demand elasticities this implies, are economically unappealing, they make perfect sense econometrically. The vehicle characteristics included in the model only capture a limited number of dimensions consumers care about. As a result the error term will include the effect of unobservables that consumers value and are willing to pay for (the ξ_j parameters in the model). Firms with price-setting power are likely to put a higher price on vehicles that have higher unobservable "quality".

This endogeneity will induce a positive correlation between price and the error term and lead to an upward bias on the price coefficient. In some applications, not taking this effect into account even leads to an upward-sloping demand curve. Expanding the number of observable characteristics will help, but it would be impossible to include every characteristic

239

consumers care about. The attractiveness of the design, a good layout of the dashboard, or reliability are only some of the characteristics that are hard to measure reliably. As a result, we will use instrumental variables to control for price setting.

We require variables that are unlikely to be correlated with unobservable aspects of a vehicle's "quality", broadly defined, but are correlated with the price. For a detailed discussion of instruments in this literature, we refer to Berry *et al.* (1995). We basically follow their insight and use as instruments the average characteristics for competing manufacturers. The observable characteristics of vehicles produced by competing firms are plausibly exogenous to the unobserved quality that consumers attach to the vehicle of one firm, while in a competitive market setting these characteristics will definitely influence the pricing decision of the firm. The discussion in Berry *et al.* (1995) includes conditions under which these instruments resemble optimal instruments. In the nested logit setting, we include two sets of instruments: average characteristics of all models produced by other firms and the same set of variables but only averaging over competing models in the same segment. Adding this second set of instruments changes the point estimates of the coefficients only marginally but improves estimation precision.

Results for the nested logit model with instrumental variables for price are in column (3) of Table 2.2. This will be the preferred set of estimates that we will use to simulate the model. The coefficient on price changes a lot. It almost doubles in absolute value—in line with the expected increase. As a result, for virtually all models in the market we find that firms are setting the price on the elastic portion of demand—in line with profit maximizing behaviour. Consumers now have a positive marginal willingness to pay for all characteristics—also as expected. The power of the engine, a standard automatic transmission[17], and vehicle size are found to be most important. Fuel efficiency and a domestic nameplate both have a positive effect on demand, but

---

[17] Using other variables to measure luxury, such as standard ABS or models explicitly marketed in luxury segments, leads to similarly positive estimates.

are not significant at usual significance levels. Finally, the nesting variable is still estimated to be large and positive, albeit not as large as with OLS. Firms in the same nest seem to resemble each other also in terms of unobservables, which is plausible.

Finally, in column (4), we also report instrumental variables estimates of the nested logit model allowing for random effects by model, to control more explicitly for unobservables[18]. These estimates provide a robustness check for the results when we control more generally for model-specific time-invariant heterogeneity. Especially, the price coefficient (α) and nesting variable (σ) are estimated extremely similar. These are the only two parameter estimates that explicitly enter the elasticity calculation (see below). As a result, own and cross-price elasticities would be very similar for the model in column (4). The coefficient on size can hardly be identified anymore, which is not surprising as this is one characteristic that manufacturers can hardly change in successive model-years.

2.3   Calculating unobserved variables

The coefficient estimates for the demand parameters in column (3) of Table 2.2 are now used to calculate the demand elasticities, marginal costs, and unobserved vehicle quality. The first two will drive the results of the trade policy simulations in the next section[19].

The demand system yields own and cross-price elasticities for all 218 vehicles for sale in Canada. A benefit of the random utility framework is that it allows a general pattern of substitution, while requiring only the estimation of a limited number of coefficients—those associated with vehicle characteristics and the degree of substitution within each nest. As mentioned earlier, for a truly flexible substitution pattern one has to allow for more random coefficients than only on the segment dummies, e.g., on price.

---

[18] Alternatively, we could estimate the model with fixed effects, but more than two years of data would be required. The random effects can be incorporated even with the limited data set we have to work with.

[19] From now on we focus on the results for 2005.

In the one-level nested logit model, the demand elasticities are:

$$\varepsilon_{jj} = -\frac{\partial q_j}{\partial p_j}\frac{p_j}{q_j} = \alpha p_j \left[\frac{1}{1-\sigma} - \frac{\sigma}{1-\sigma}\frac{q_j}{Q_g} - \frac{q_j}{L} \right] \qquad j \in \text{nest } g$$

$$\varepsilon_{jk} = \frac{\partial q_k}{\partial p_j}\frac{p_j}{q_k} = \alpha p_j \left[\frac{\sigma}{1-\sigma}\frac{q_j}{Q_g} - \frac{q_j}{L} \right] \qquad \text{both } j \text{ and } k \in \text{nest } g$$

$$\varepsilon_{jk'} = \frac{\partial q_{k'}}{\partial p_j}\frac{p_j}{q_{k'}} = \alpha p_j \left[\frac{q_j}{L} \right] \qquad j \in \text{nest } g, k' \in \text{nest } g' \neq g$$

Table 2.3 Own and cross-price elasticities for a select number of models

	Civic	Mazda3	Pursuit	Elantra	Golf	Escape	CR-V	Santa Fe	Equinox
Civic	-2.499	0.254	0.254	0.254	0.254	0.009	0.009	0.009	0.009
Mazda3	0.195	-2.574	0.195	0.195	0.195	0.007	0.007	0.007	0.007
Pursuit	0.137	0.137	-2.570	0.137	0.137	0.005	0.005	0.005	0.005
Elantra	0.062	0.062	0.062	-2.487	0.062	0.002	0.002	0.002	0.002
Golf	0.029	0.029	0.029	0.029	-3.121	0.001	0.001	0.001	0.001
Escape	0.004	0.004	0.004	0.004	0.004	-3.680	0.229	0.229	0.229
CR-V	0.004	0.004	0.004	0.004	0.004	0.210	-4.584	0.210	0.210
Santa Fe	0.002	0.002	0.002	0.002	0.002	0.121	0.121	-3.448	0.121
Equinox	0.003	0.003	0.003	0.003	0.003	0.152	0.152	0.152	-4.372

Note: The statistics indicate the demand elasticity of the model in column for price changes of the model in the column. Own price elasticities are on the diagonal.

We calculate the own and cross-price elasticities between all models—$J*(J+1)/2$ elasticities (23871 elasticities in 2005)—as they are used to uncover the marginal costs the model implies. Table 2.3 lists the own and cross-price elasticities for a select number of vehicles from the two largest Canadian market segments. For each region of the world[20]—Canada, U.S. and

---

[20] These five regions will be considered separately in the trade policy simulations below.

Mexico, South Korea, Japan, and the E.U.—we include the best-selling vehicle. The first five models are from the "small car" segment, which combines the lower small, upper small and small specialty cars according to the market segmentation in Ward's Automotive Yearbook. The next four models are "small SUVs", a nest that combines the small and middle SUVs and small and middle cross-utility vehicles (car-based SUVs) segment[21].

A crucial—and admittedly undesirable feature—of the nested logit demand model is that within each nest the absolute value of the own-price demand elasticity is an increasing function of a model's price. This follows directly from the functional form for demand. Within each segment, all vehicles share the same demand curve, except for the random individual-model specific logit error draw. As a result, more expensive models will be priced higher up on the demand curve, where consumers are more elastic.

The cross-model elasticity of substitution is much higher for models in the same nest, driven by the high estimate for σ, and the elasticity of substitution between models in all other nests is the same. Ideally, we would let the added substitutability within each nest vary and estimate seven distinct σ parameters. Brenkers and Verboven (2006) illustrate that with such added flexibility substitution parameters in more expensive segments tend to be lower and own-price elasticities do not have to rise with price. However, given that we only have two years of data available, several of the estimated σ parameters were estimated higher than 1, although not significantly so, which is inconsistent with a well-behaved demand system. Therefore, we were forced to impose similarity of the σ parameter in each nest and as a result demand elasticities increase with price throughout. A factor that exacerbates this tendency is that more expensive segments in Canada tend to be more crowded, increasing demand elasticities further. The elasticity formula

[21] In the market segmentation followed in DesRosiers publications for Canada, "small cars" refers to subcompacts, compacts, and smaller sport cars, "small SUVs" would refer to compact SUVs, SUVs, and intermediate SUVs.

clearly indicates that own-price elasticity is increased if a vehicle has only a low market share within its segment.

The model does perform well in predicting different cross-price elasticities for models that are in the same segment and those that are not. For example, a 10% price increase for the Honda Civic, the most popular model, raises the expected sales of all other small cars by 2.54%. The effect on models in all other segments, including the outside good (i.e. second hand cars), is much smaller, a 0.9% sales increase. A similar price increase for the Mazda3 leads to only a 1.95% sales increase for other models in the small car segment. The difference is explained by the fact that the Mazda has only ¾ of the sales of the Civic, so given that the own-price elasticity is similar (-2.574 versus -2.499) the number of lost Mazda sales that spill over to competitors is proportionally lower.

In the consumer's random utility function is an unobservable model-specific characteristic (ξ_j) that directly enters the (normalized) market share function linearly—in our estimation it becomes the error term of the regression. This is the next quantity that can be calculated from the fitted demand model. While it will not play an independent role in the results, since results only depend on the joint effect of the entire part of the utility function that is common to all individuals' valuation (δ_j), it provides a useful check for plausibility of the model estimates. The average ξ_j is zero for the entire sample, but to aid comparability we normalize it to zero by segment. Vehicles with positive ξ_j have a higher demand than one would predict based on the observable characteristics. As such, it measures the unobservable "quality" of the vehicle.

Table 2.4 contains the name, production region, sales, and price of the same select group of vehicles that were included in Table 2.3. To give some idea about the relative position in their segment, their sales rank is also included[22]. The next column lists the unobserved "quality" of the vehicle, which by and large corresponds to our priors. The Honda Civic and Mazda3 record

---

[22] For the small SUVs, I did not include the sales leader, the Pontiac Montana SV6 because it was newly introduced in 2005.

positive unobserved quality, indicating that their sales are higher than one would expect based on the observable characteristics included in the demand estimation. In the case of the Civic, high resale value springs to mind, and for the Mazda3 the original new design can be noted. The Pontiac Pursuit, on the other hand, has a large negative quality. It indicates that the model would be expected to sell in much higher numbers, especially given its low price, strong engine, and large size.

Table 2.4 Calculated unobservables for a select number of models

	produced	Sales	rank in segment	price	unobserved "quality"	marginal cost	mark-up
Honda Civic	Canada	63676	1	16200	0.523	9697	0.401
Mazda3	Japan	48576	2	16295	0.321	9616	0.410
Pontiac Pursuit	U.S. & Mex.	34852	4	15925	-0.823	8513	0.465
Hyundai Elantra	South Korea	16711	11	14994	-0.010	8655	0.423
Volkswagen Golf	E.U.	6258	21	18530	0.095	12428	0.329
Ford Escape	U.S. & Mex.	21466	2	22995	0.324	16239	0.294
Honda CR-V	Japan	16019	3	28200	0.266	21872	0.224
Hyundai Santa Fe	South Korea	12383	5	20995	-0.155	14589	0.305
Chevrolet Equinox	Canada	12291	6	26614	0.119	19042	0.285

Note that in absence of actual transaction prices we use MSRP as the price. The domestic manufacturers tend to discount their selling price more than imports. As a result, the unobservable quality that would be imputed for the Pontiac Pursuit if actual prices would be available would likely be even lower. It would be preferable to use transaction prices instead of MSRP, but data limitations make this impossible[23]. Another notable pattern is that both Korean entries in Table 2.4 have a below average imputed unobservable quality. Sales for these models are lower

---

[23] Note that J.D. Power collects transaction information for the U.S., but the cost of these data far exceeds the budget for this study.

than expected, especially given their attractive low price. This coincides with the general perception that although the quality of the Korean cars and SUVs has improved spectacularly over the last decade, they have not closed the gap with their Western or Japanese competitors entirely. Because we only included relatively successful vehicles in Table 2.4, average quality tends to be high. Vehicles ranked much lower in their segment tend to have lower imputed quality as well—partly as explanation of their poor sales performance.

Once we add a first order condition for price setting to the estimates of the demand system, we can uncover what marginal costs for each vehicle have to be to rationalize the observed prices. We assume that firms compete in prices and that observed prices are at equilibrium in a differentiated products (Bertrand) pricing game. Moreover, firms are explicitly modeled as multi-product firms, taking into account the effect of the price of each model on all the other models they own[24]. For a derivation of the first order condition, we refer the interested reader to Berry (1994) or Berry, Levinsohn, and Pakes (1995). The imputed marginal costs for a select group of vehicles are in the second last column of Table 2.4.

Finally, the mark-up on each vehicle, defined as $\dfrac{p_j - mc_j}{p_j}$

is in the last column. Note that we have explicitly incorporated multi-product behaviour by the firms. The effects can be seen by comparing the mark-up on the Pontiac Pursuit and the Hyundai Elantra. While the marginal costs of both vehicles are similar, GM chooses to put a much higher mark-up on the Pontiac. The reason is simply that 9 of the 32 models in the small car segment are owned[25] by GM and it takes into account that lowering the Pontiac's price will to a large extent merely cannibalize the sales

of its other offerings in the segment. The substitution patterns in Table 2.3 illustrate clearly that the bulk of cross-model substitution happens within each segment. The same reasoning explains the similar mark-up between the Chevrolet Equinox and the Hyundai Santa Fe in the small SUV segment, even though the estimated marginal cost for the Chevy is $4,500 higher.

The much smaller mark-ups for the small SUV segment versus the small car segment results from the higher average price for SUVs combined with the restrictive functional form assumption for demand. Given that the substitution patterns within each segment dominate and that prices are relatively similar within each segment, the impact of this on the trade policy simulations is likely to be second order as cross-product substitutions are not affected by this.

We now have all the ingredients—a demand system, imputed marginal costs for each model, and a market equilibrium assumption—to turn to the counterfactual policy experiments.

2.4   Simulating trade policy changes

In this section, we rely on the previously discussed results to perform four counterfactual policy simulations. We look at the impact of four trade policy changes on a number of important economic variables. The policy changes for Canada that we consider are:

- FTA (only) with South Korea
- FTA (only) with Japan
- FTA (only) with the E.U.
- Unilateral abolition of the Canadian import tariff on final vehicles

Currently, Canada imposes a 6.1% import duty on finished vehicles. In each of these four scenarios we will investigate how the market equilibrium would look differently if vehicles imported from one or more countries would be exempt from the import duty. A number of caveats are in order before we turn to the discussion of the results:

1. We do not consider domestic content requirements in this exercise. It is likely that any FTA agreement would specify domestic content rules, much like those in force under

NAFTA. We simply assume that all relevant firms would (costlessly) be able to satisfy those rules[26].

2. One should not interpret the results as a prediction of the likely future effects of such trade policy changes. We calculate what the market equilibrium would have looked like in 2005 if an alternative trade regime would have been in effect.

3. The results do take into account responses of only indirectly affected firms. The competitive situation and hence the optimal prices for domestically produced vehicles will differ if one or more importers are suddenly exempt from import duties. We let all market participants adjust to the new situation. Hence, our results should be interpreted as long-term effects.

4. We only vary the marginal costs of firms for which the import regime changes—as they do not have to pay duties anymore. In order to impute the marginal costs that rationalize the observed price vector, we had to assume marginal costs are constant, i.e. do not vary with output[27]. As such, the only thing that changes for domestic producers is the degree of competition.

5. The effects of the newly opened Hyundai plant in Alabama and the Toyota plants under construction in Texas and Baja California are not incorporated yet into this analysis.

6. With our model we are able to analyze the domestic Canadian market. We will study the impact of trade policy on sales, pro-

---

[26] As noted earlier, in the last two policy simulations we do not modify the duty treatment of the vehicles assembled by BMW or Mercedes-Benz in the U.S. Currently these vehicles do not satisfy the NAFTA domestic content requirements and incur duties when imported into Canada. Under an FTA with Europe or under unilateral free trade by Canada it would be reasonable to assume these vehicles would also be exempt from duty (as their joint E.U./NA domestic content will far exceed any plausible threshold). It would have been too time consuming to adjust our simulation programs to take this into account. Note that only 5000 vehicles annually are affected by this shortcoming, a mere 0.3% of the Canadian market.

[27] This assumption is made throughout in the literature. Relaxing it would directly affect all firms' first order conditions and severely complicate the calculation of a new equilibrium.

duction, imports, prices, mark-ups, profits, consumer surplus, and tariff revenue. We will break down the impact by the origin of production—produced in Canada or imported from one of the four other regions. However, we do not look at total Canadian production. Demand in other countries is unlikely to be affected in any important way by a change in Canadian import tariffs. As a result, Canadian production for export is assumed to remain unchanged. Furthermore, we cannot discuss the impact of an FTA with China or Mercosur, because that would be mere speculation at this point—how to know the elasticity of substitution between a Dodge Caravan and a not-yet-introduced Chinese-made vehicle?[28]

With these caveats in mind, we now turn to the results from the trade policy simulations. The actual and predicted levels of all economic variables are in Table 2.5. The four different policy changes are reported in the different columns. Table 2.6 contains the same results, but shows all effects as percentage changes relative to the 2005 baseline case. First, we discuss our calculations by introducing the results for the actually observed market equilibrium in 2005.

2.4.1   *The baseline case: 2005 Canadian automobile market*

The actual quantities of all the relevant economic statistics for 2005 are in the first column of Table 2.5. The average quantity-weighed price was just over $25,000. The average mark-up, again weighted by sales, was 31.1% which implies that the average marginal cost was $19,124. Note that this marginal cost excludes all fixed costs involved in making, marketing, and selling a vehicle: designing the vehicle, building and maintaining an assembly plant, retooling all capital equipment, all advertising and marketing expenses that are independent of the actual number of cars sold, fixed costs of maintaining a dealership network, etc.

---

[28] In Section 4, when we study automotive components, those regions will be studied.

The average price of vehicles produced in Canada is $822 below the national average, while the average American- or Mexican-made vehicle is $591 more expensive than the national average. Not surprisingly, the average European import is much more expensive—at $41,728—and the average Korean import much cheaper—at $17,678—while Japanese vehicles most closely resemble Canadian vehicles. In line with the earlier discussion, we find again that more expensive vehicles are associated with lower mark-ups. Note that throughout we will use the term Canadian vehicles for vehicles produced in Canada, including foreign nameplates such as the Honda Civic or Lexus RX220. Korean vehicles, on the other hand, will include vehicles badged by Hyundai and Kia, but also some Chevrolets and Suzukis. In the same spirit, Canadian profits are meant to indicate all variable profits made on vehicles produced in Canada, irrespective of the owner.

Total sales in Canada in 2005 amounted to almost 1.6 million vehicles, cars and light trucks combined; more than ¼ of these vehicles were assembled domestically[29]. Of course, the vast majority of cars assembled in Canada are exported, but as mentioned earlier, we assume Canadian exports are unaffected by trade policy changes and do not discuss them further. Canadian imports total 1.17 million and just over 2/3 of these come from the U.S. or Mex-

---

[29] Total Canadian sales of models that are produced in a Canadian assembly plant (and possibly in other plants as well) adds up to 401,292 units for the 2005 model year. Note that total production of these models is much higher as the majority of output is exported. Note also that actual domestic sourcing of Canadian-made vehicles is bound to be lower as firms produce some of their highest volume vehicles in a second assembly plant in the U.S. As actual configurations produced differ between plants, some Canadian demand for a vehicle produced domestically will be filled by U.S. plants. For example, in 2005 DaimlerChrysler sold 216,857 vehicles in Canada, 34,979 of these were produced locally. Total sales in Canada of the Dodge Caravan, produced in Windsor and St. Louis, exceeded 60,000. Similar problems exist for GM (Canadian demand of the Chevrolet Silverado and GMC Sierra, produced in Oshawa, exceeds their "total Canadian production for sale in Canada". For Honda (Civic) and Toyota (Corolla), the problem exists as well, but it is not as large. In absence of information at the model level of the final destination of vehicles, we are forced to use the definition of Canadian production we adopted.

ico, entering the country duty-free under NAFTA. The market share of cars made in Japan is 11.3%, which translates into a Japanese import share of 15.2%. For South Korea, the comparable statistics are 8.6% and 11.5% and, for the E.U., market and import shares are 4.5% and 6.0% respectively. Note that these shares differ from those in Table 2.1 as the model characteristics in that table are not weighted by sales volumes[30].

Converting the implied consumer surplus into a dollar amount using the estimated price coefficient yields a surplus of $33.8 billion—or an average of almost $29,000 per sold vehicle. This is the aggregate utility value over and above the sales price consumers attach to their new vehicle purchases. This implausibly high estimate is due to the fact that consumers get vehicle-specific draws in their utility function and, as a result, people tend to buy cars that give them a high utility level for factors mostly unexplained by the model. While this is a major problem investigating the introduction of new goods, see Petrin (2002), in the current application we do not change the range of models for sale in the market. While the level of the surplus is likely to be unreliably estimated, we will only look at changes. Aggregate variable firm profits are on the order of $11 billion, 26.7% of which are earned on vehicles made in Canada—approximately in line with the Canadian production share. Note, once again, that these are variable profits and that they include all the fixed costs firms incur. They are entirely incomparable to the accounting profits that firms have to report. Given that fixed costs are, by definition, fixed, we can still use the profit measure to get a reliable estimate of how trade policies will affect firms' profitability.

---

[30] While only 10% of the observations are models produced in Canada, they represent 25% of sales.

Table 2.5 Trade policy simulations: levels (model year 2005: Sept. 2004 – Aug. 2005)

	Actual situation in 2005	FTA with: South Korea	FTA with: Japan	FTA with: E.U.	Unilateral elimination of Canadian tariff
Aggregate effects on:					
Price (average)	$25,134	$25,045	$25,066	$25,372	$25,210
Mark-up (average)	31.3%	31.5%	31.4%	31.2%	31.5%
Demand	1,574,635	1,578,561	1,583,037	1,581,758	1,593,770
Canadian production	401,292	399,155	397,500	398,327	392,624
Imports (NA + ROW)	1,173,343	1,179,406	1,185,537	1,183,431	1,201,146
Consumer surplus (mil.)	33,819	33,914	34,022	33,991	34,283
Firm variable profits (m)	11,034	11,053	11,071	11,078	11,131
Firm profits in Canada	2,948	2,930	2,920	2,926	2,881
Tariff revenue	426	333	235	270	
Domestic welfare (mil.)	37,193	37,177	37,177	37,187	37,164
Effects, broken down:					
Prices					
– Canada	$24,312	$24,314	$24,304	$24,256	$24,253
– U.S. & Mexico	$25,725	$25,732	$25,673	$25,598	$25,559
– South Korea	$17,678	$17,343	$17,648	$17,674	$17,309
– Japan	$23,505	$23,504	$23,603	$23,262	$23,343
– E.U.	$41,728	$41,774	$41,546	$44,016	$43,913
Mark-ups					
– Canada	32.4%	32.4%	32.4%	32.4%	32.4%
– U.S. & Mexico	29.9%	29.9%	29.9%	30.0%	30.0%
– South Korea	40.7%	41.7%	40.7%	40.7%	41.6%
– Japan	33.1%	33.1%	33.3%	33.3%	33.4%
– E.U.	18.9%	18.8%	18.9%	18.8%	18.8%
Quantities					
Canada (production)	401,292	399,155	397,500	398,327	392,624
U.S. & Mexico (import)	789,553	784,260	779,134	782,880	767,776
South Korea (import)	135,378	148,538	133,259	134,913	145,769
Japan (import)	178,319	176,753	205,255	175,698	200,315
E.U. (import)	70,093	69,855	67,889	89,940	87,285

Finally, we can also calculate the government's tariff revenue from imported vehicles. The model does not provide us with the import value of the vehicle that we can use to calculate the duty on. In absence of any other plausible magnitude, we use the estimated marginal cost for each vehicle as base to calculate duties (on average, across all vehicles, marginal cost is 70% of the final consumer price). While this excludes some of the fixed costs likely to be subject to tariffs, it includes any costs incurred in the distribution channels which should be excluded. Overall, we are not likely to misestimate tariff revenue by much and in addition, we are mostly concerned with changes over time. In 2005, our assumption leads to Canadian tariff revenue of $426 million, or just above $1000 per imported vehicle on average—note that vehicles imported from the U.S. or Mexico are excluded from duties.

Our measure of domestic welfare in the final goods sector of the industry is the sum of consumer surplus, profits earned on vehicles assembled in Canadian plants, and the government's tariff revenue.

2.4.2   FTA with South Korea

In the second column of Table 2.5, the relevant statistics are reported calculated from a new industry equilibrium where Korean imports are not subject to the 6.1% import tariff anymore. The first column of Table 2.6 contains the same results, expressed as changes from 2005. For the discussion, we will focus on Table 2.6[31].

---

[31] These statistics are calculated by computing a new price equilibrium from the vector of first order conditions for all firms. Bresnahan (1987) contains a very clear discussion of the derivation of the first order condition for multi-product firms. The marginal costs for Korean-made vehicles are lowered by $1/1.061$ and using a contraction mapping the new price vector is calculated. Note that all elasticities and cross-price elasticities enter the first order conditions and influence the calculated price change. From the estimated demand system we can then calculate all new quantities, profits, trade flows, profits and consumer surplus.

Table 2.6 Trade policy simulations: changes

| | FTA with: | | | Unilateral elimination of Canadian tariff |
	South Korea	Japan	E.U.	
Aggregate effects on:				
Price (average)	-0.35%	-0.27%	0.95%	0.30%
Mark-up (average)	0.16%	0.07%	-0.09%	0.11%
Demand	0.25%	0.53%	0.45%	1.22%
Canadian production	-0.53%	-0.94%	-0.74%	-2.16%
Imports (NA + ROW)	0.52%	1.04%	0.86%	2.37%
Consumer surplus	0.28%	0.60%	0.51%	1.37%
Firm profits	0.17%	0.33%	0.40%	0.88%
Firm profits in Canada	-0.61%	-0.96%	-0.77%	-2.29%
Tariff revenue	-21.83%	-44.84%	-36.62%	-100.00%
Domestic welfare	-0.04%	-0.04%	-0.02%	-0.08%
Effects, broken down:				
Prices				
– Canada	0.01%	-0.03%	-0.23%	-0.24%
– U.S. & Mexico	0.03%	-0.20%	-0.49%	-0.65%
– South Korea	-1.90%	-0.17%	-0.02%	-2.09%
– Japan	0.00%	0.42%	-1.03%	-0.69%
– E.U.	0.11%	-0.44%	5.48%	5.24%
Mark-ups (percentage point change)				
– Canada	-0.03%	-0.01%	0.03%	-0.01%
– U.S. & Mexico	-0.02%	0.00%	0.07%	0.05%
– South Korea	0.96%	0.00%	-0.01%	0.94%
– Japan	-0.02%	0.17%	0.18%	0.34%
– E.U.	-0.03%	0.02%	-0.09%	-0.11%
Quantities				
Production — Canada	-0.53%	-0.94%	-0.74%	-2.16%
Imports — U.S. & Mexico	-0.67%	-1.32%	-0.85%	-2.76%
Imports — South Korea	9.72%	-1.57%	-0.34%	7.68%
Imports — Japan	-0.88%	15.11%	-1.47%	12.34%
Imports — E.U.	-0.34%	-3.14%	28.32%	24.53%

254

The average price in the Canadian market is predicted to be 0.4% lower under an FTA with South Korea than the actual price observed in 2005. This is the combined effect of four influences. We will discuss each of these tendencies in detail for the FTA with Korea, but the same factors will operate in all other trade policy simulations. In the different policy scenarios discussed in the following sub-sections, the relative importance of each effect will vary substantially.

First, as a result of the import duty exemption, South Korean producers have a lower marginal cost which, ceteris paribus, lowers the market price. If their mark-ups would have been unchanged and if there were no response from competitors, all Korean prices would have been reduced by 5.75% (1-1/1.061). The model predicts the average Korean price to decline by only 1.9%, so more factors are at work. The pass-through to consumers of the tariff elimination was clearly less than 100%.

The second effect, which is directly within the Korean firms' control, is that, with the new marginal costs, optimal price-cost margins change. In particular, given that costs are lower, without changing mark-ups Korean vehicles would be priced at a lower point on the demand curve, where the elasticity of substitution is lower, and the optimal response would be to increase prices. The results indicate that the average mark-up did increase by 1%. We should point out that the size of this effect is likely to be overestimated because our functional form of demand imputes a very low demand elasticity for cheaper cars, a defining feature of many Korean vehicles. Furthermore, while the sole Korean firm is obviously most affected by this policy change—Hyundai imports 14 models that it assembles in Korea into Canada—GM's Daewoo subsidiary also exports 6 models to Canada, two of which are badged as Suzukis, one as a Pontiac and three as Chevrolets. Almost 1/3 of Korean imports in Canada are GM products.

Third, competitors will react to the Korean price cuts—the net effect of the lower marginal cost and the higher mark-up is to lower prices. The results indicate that the competitive responses of competitors are limited. Statistics in Table 2.7 indicate that the raw average price change of vehicles produced

255

in Korea is -3.57%, while the price drop is several orders of magnitude smaller, between -0.01% and -0.03% for goods produced elsewhere. Note, once again that the price response of other producers will in turn lead to successively smaller price responses of Korean firms, etc. The statistics in Table 2.7 are the result of the eventual convergence of all these price responses, where no firm has any incentive to change its price anymore.

European producers have very little overlap with Koreans, most of their vehicles are in different segments, and they have the lowest price response. Several vehicles made in Canada, the first line in Table 2.7, are in segments where Korean vehicles are important, lower cars and lower SUVs and we find a larger response for Canadian-made vehicles. Disaggregating the price changes in Table 2.7 further (numbers not reported in the table) would reveal that Canadian-made vehicles in the small car segment see a 0.07% price drop, while those in the luxury car segment become only 0.02% cheaper. Similarly, averaged over all non-Korean producers, small SUVs become 0.05% cheaper while there is no noticeable price change in the upper SUV segment—where no Korean-made cars are sold.

Table 2.7 Average price change in response to trade policy change

Vehicles produced in:	FTA with:			Unilateral elimination of Canadian tariff
	South Korea	Japan	E.U.	
Canada	-0.03%	-0.02%	-0.06%	-0.11%
U.S. & Mexico	-0.02%	-0.05%	-0.02%	-0.08%
South Korea	-3.57%	-0.08%	-0.02%	-3.67%
Japan	-0.02%	-4.44%	-0.07%	-4.54%
E.U.	-0.01%	-0.06%	-4.81%	-4.87%

Fourth, composition effects cannot be ignored. In Table 2.6, average price changes are minimal for all other regions, but looking at further digits reveals that they are positive. As the results in Table 2.7 clearly indicate, this does not imply that firms actually increase their prices—in fact the price of every single vehicle sold in Canada declines with the FTA. Rather, it

implies that the composition of goods sold changes. Given that Korean products—which tend to be priced at the low end of the market—are even more competitive after the trade policy change, other producers lose sales there, which changes the weight on their average price towards more expensive vehicles.

Similarly, the 5.75% decrease in Korean marginal costs combined with the 1% increase in the mark-up does not translate into 4.75% lower prices because the composition of sales also changes for Korean firms. This is the result of two factors that lead to a higher relative weight on more expensive vehicles also for Korean producers. First, given that the demand elasticity is estimated to be increasing in price, a much greater fraction of the tariff savings are passed along to consumers of more expensive vehicles. It improves the competitive position of Korean vehicles much more in more expensive segments (middle cars and lower SUVs). Second, Korean firms are not as well represented in these upper segments, so their lower prices are less likely to lead them cannibalizing their own sales. In the lower SUV segment only 10% of the models are produced in Korea, as opposed to 30% of lower car models. Both factors lead to higher sales increases for more expensive Korean vehicles, which increases their average price.

Given this elaborate discussion, the rest of the results should be straightforward:

- Average prices fall slightly, which is mainly driven by an imperfect pass-through of the tariff reduction on Korean vehicles and to a lesser extent the result of competitive responses by other producers.

- Average mark-ups increase for Korean firms, mainly as a result of their lower marginal cost. Foreign firms lower their mark-ups slightly, both as a competitive response to the Koreans and as a compositional effect as their sales become more heavily weighted towards expensive vehicles.

- Aggregate vehicle sales increase, not surprisingly, as the average price of every vehicle sold in Canada declines. The pattern follows the mark-ups. The magnitude of the increase is lower than the price increase (even though virtually all ve-

257

hicles are priced on the elastic portion of demand), because mark-ups increase as well.

- Korean imports increase, while all other regions—including production in Canada—lose. We conjecture that the current demand system—with elasticities being uniformly increasing in price—underestimates the impact of the Korean output response. The Korean import response is sufficiently large that net Canadian imports increase for sure.

- Lower prices lead to a higher consumer surplus, but less profit is made on vehicles produced in Canada—providing opposite effects on aggregate welfare.

- Tariff revenue for the government is reduced by almost 22%.

- If we look at aggregate welfare in dollars terms, Table 2.5, we find that consumer surplus increases by $95m, made-in-Canada profits fall by $18m, and government revenue falls by $93m, for a net Canadian loss of $16m or a mere 0.4% of the welfare generated in this industry. Two caveats go with this finding. First, a demand system that estimates a higher demand elasticity for vehicles made in Korean—which seems plausible—would increase the benefits. Second, the loss in firm profits will to some extent accrue to the foreign owners of the Canadian plants (U.S. and Japanese corporations)—although part of the increase in variable profits might be captured by the workforce. This might lead one to discount the profit loss in Canadian welfare calculations.

2.4.3 FTA with Japan

The gist of the analysis associated with a Japanese FTA is similar to the analysis in the preceding sub-section. The demand elasticities in Table 2.8 preview the effects one can expect. Even if we limit attention to cars—the pattern on light trucks which are on average more expensive would be similar—Korean firms have a median demand elasticity significantly below the median for all other production regions, especially below Japanese or European imports. The last column of Table 2.8 contains these demand elasticities for all cars. As mentioned

earlier this is a combination of the functional form of our demand system as well with a crowded product space in more expensive segments of the market.

Importantly, the large difference in the last column is for the most part a compositional effect. Within each market segment the differences between the regions are much smaller. Korean cars simply tend to be positioned predominantly in the small car segment, where demand elasticities are estimated to be low for all producers. This fact will work its way through the entire analysis.

One notable effect is that even though Japanese firms lower their prices on every single vehicle in every segment—90% of the Japanese price reductions range between 3.6% and 5.1%, indeed very close to complete pass-through—the composition effects are so strong that the average sales-weighted Japanese price ends up 0.42% higher under an FTA with Japan. This is largely the result of very expensive Lexus, Acura, and Infiniti products that see relatively large sales increases and pull up the average Japanese price. Given that these models are priced at the very elastic point of the demand curve, pass-through of the tariff savings is almost perfect, while at the same time consumers are estimated to be very price responsive.

Table 2.8 Median own-price demand elasticity for all car segments by car segment:

Vehicles produced in:	Small	Medium & large	Luxury	All cars
Canada	-2.47	-3.74	-3.29	-3.44
U.S. & Mexico	-2.60	-3.94	-8.14	-4.01
South Korea	-2.05	-4.56		-2.52
Japan	-2.57	-4.48	-7.04	-5.00
E.U.	-3.04	-4.75	-9.46	-6.99

Mark-ups are estimated to change less than with the Korean FTA, which is the result of a much smaller share of Japanese vehicles in the lowest price segments. The best-selling Toyota,

Honda, and Nissan small cars are all produced in North America. Given that mark-ups are estimated to be lower in the more expensive segments, we find less of a response by Japanese firms as well as by competitors. Notably, the largest effect is for European producers which compete with Japan-made cars in all luxury segments.

Because the estimated pass-through of the tariff savings is higher than in the Korean FTA case and because the estimated demand elasticity for Japanese products is higher on average, the estimated import increase of Japanese cars and light trucks is higher, at 15.1%. This increase works its way into higher total imports in Canada, 1.04% higher, and lower domestic production, a change of -0.94%.

While the average price drop is lower under an FTA with Japan than with Korea, -0.27% versus -0.35%, to a large extent this is caused by consumers trading up and purchasing more expensive Japanese imports. The increase in consumer surplus is twice as high as in the previous analysis, +0.60% versus +0.28%.

While 2005 imports of Japanese vehicles were only 32% higher than Korean imports, the average value of these vehicles was much higher. As a result, the cost of the FTA in lost tariff revenue for the Canadian government is estimated to be more than twice as high as in the previous sub-section, -44.8% versus -21.8%. As a result of this final factor, and in spite of the robust consumer gains with a Japanese FTA, overall Canadian welfare is estimated to be lowered by exactly the same amount as in the preceding analysis, -0.04% or $15m. However, the distribution of that amount is noticeably different. Consumers would gain $203m, more than twice as much, while the bulk of the loss would fall on the government. Of course, indirectly this burden falls on the taxpayers, approximately 80% of which bought the outside good (i.e. not a new car) in 2005.

2.4.4   FTA with the E.U.

Finally, the higher demand elasticity of the median European car in every segment, see Table 2.8, leads to qualitatively similar results as in the Japanese FTA case, but with even stronger

compositional effects. For example, the average price is predicted to increase, as the extremely expensive European vehicles gain market share. The same compositional effects, lead to a lower weighted-average European mark-up, as percentage mark-ups are estimated to be lower for more expensive vehicles. The intermediately-priced Japanese producers, on the other hand, see sales of their most expensive vehicles decline, which raises their average mark-up.

The price responses are rather impressive. The average Canadian price is estimated to increase by almost a full percentage point, even though the average mark-up goes down. Consumers trade up to more expensive vehicles very aggressively, which leads to an increase in consumer surplus almost as large as under the Japanese FTA—even though European imports numbered less than half of Japanese imports in 2005. This surge in expensive car purchases also boosts firm profits, which on average rise by 0.40%, although lower domestic production hurts Canadian producers. As before, the loss of tariff revenue on the expensive European imports is estimated to set Canadian tariff revenues back a full 36.6%, a revenue loss almost twice as high as under the Korean FTA even though Korean imports, in units, were almost double European imports in 2005. When all is said and done aggregate welfare hardly budges, falling by $6m or about 20c per Canadian.

We do not discuss these results at length because they depend crucially on the high demand elasticity for expensive vehicles. 30 of the 38 European imports are in the luxury car or large and luxury SUV segments. Even the three European entries in the small car segment are among the 10 most expensive vehicles in that segment. While only 17% of models sold in Canada are assembled in Europe, 56% of the luxury car segment entries are.

Moreover, the 38 European imports are sold by 6 different firms while the 20 Korean imports are sold by only 2 firms. As a result, in the case of a European FTA, multi-product considerations are not holding firms back from lowering their prices.

Figure 2.1 Price responses in the three FTA scenarios
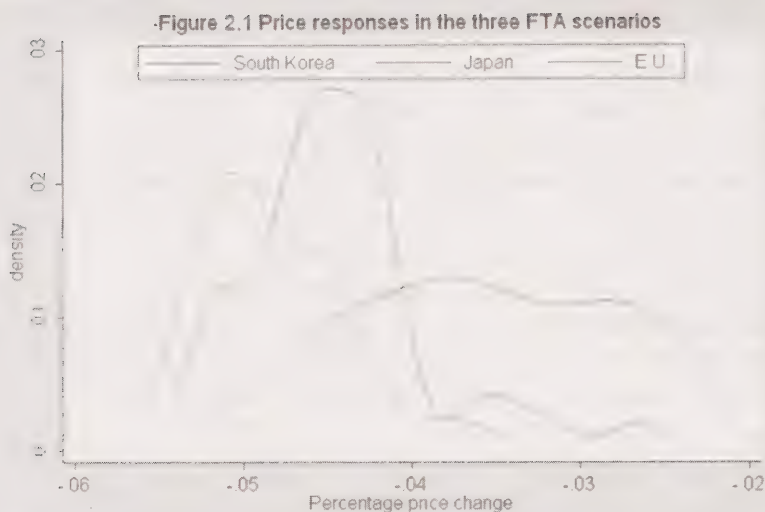
The effects of this very different market presence and ownership structure for firms from the different regions shows up directly in the price adjustments to the different FTAs. The lines in Figure 2.1 plot a smoothed histogram for the distribution of percentage price changes under the three different scenarios only for the models that gain a direct advantage of the policy change, i.e. the price changes for Korean-made vehicles are plotted only in the case of the Korean FTA, and similarly for the other two regions. The green line represents the distribution of price responses for European-made cars under an FTA between Canada and the E.U. Clearly price changes are concentrated around -5.1%, very close to the 5.75% that would indicate complete pass-through of tariff changes. In contrast, the red line for vehicles made in Japan shows many more intermediate price changes, around -4.3%, while the blue line for Korea indicates that for many of those models price reductions are less than half of the tariff reduction.

While the extent of substitution between imports and domestic production and between the imports of the different countries is likely to be robust to the other specifications of the demand system, this difference in average price responsiveness

262

by region hinges crucially on the average difference in demand elasticity. It is not impossible that consumers purchasing the expensive vehicles are indeed as price sensitive as the demand model predicts, but to increase confidence in the results, we would like to see how high the α coefficient on price would be estimated when price changes over time are used to identify the coefficient, rather than an identification solely from the cross-section of vehicles as is currently the case. Also, a more general demand model should ideally either incorporate a random parameter on the price or different nesting parameters by segment. Unfortunately, incorporating these changes would require much more data and take a lot of time.

2.4.5   Unilateral elimination of the Canadian import tariff

Finally, a unilateral elimination of the import tariff by Canada, results are in the final columns of Tables 2.5 and 2.6, is predicted to lead to the largest drop in aggregate welfare of the four trade policy experiments. One might be surprised by this finding, as the common economic wisdom predicts that free trade is good for welfare, or not? One should not forget that this is a concentrated industry with differentiated goods and firms are expected to have a lot of market power. Moreover, distributional effects between consumers, domestic and foreign profits, and the government are crucial.

In particular, consumers are estimated to gain the equivalent of $464m in consumer surplus, while Canadian producers are expected to lose $67m on their domestically produced vehicles –relative to the 2005 baseline. However, these same firms are also importing a lot of vehicles, and on average the worldwide firm profits are predicted to increase by $97m.

The results in Table 2.9 break the aggregate profits down by firm. Firms that rely more on imports are likely to gain most. Most prominent are Hyundai, BMW, Nissan, and Ford. Note, however, that the new Hyundai plant in Alabama will lower the expected benefits that Hyundai can hope to achieve from tariff elimination. Only firms that produce a large fraction of their Canadian sales domestically, especially GM and to a smaller

263

extent also Honda and DaimlerChrysler, stand to lose from the Canadian elimination of the import tariff.

Table 2.9 Change in profits without the Canadian import duty (by firm)

Corporate group	2005 profit	Change in profit without import duty	
	(million $)	(million $)	(%)
GM	$3,601	-$60	-1.67%
Ford	$2,136	$22	1.04%
DaimlerChrysler	$1,625	-$4	-0.25%
Toyota	$1,150	$15	1.33%
Honda	$962	-$11	-1.10%
Hyundai	$585	$53	9.00%
Nissan	$456	$24	5.35%
Volkswagen	$219	$13	6.17%
BMW	$140	$25	18.04%
Subaru	$93	$8	8.92%
Mitsubishi	$58	$3	4.90%
Porsche	$12	$8	64.23%
Total	$11,034	$97	0.88%

Note: The corporate groups include partially owned subsidiaries: GM includes Suzuki and Ford includes Mazda

So while consumers would gain from such a trade policy and most of the firms would as well, the Canadian government would lose $426m in tariff revenue, or approximately $24 per labor market participant, which is not negligible. On the other hand, the higher price for the average vehicle combined with higher demand would increase sales tax. Additional GST revenues would run to $42m for the federal government and a similar amount for the provinces. These numbers are not included in the welfare calculations because a gain for the government would be a loss to consumers.

264

At the same time Canadian production is estimated to decline by 8,668 units annually. While this is not nearly enough to noticeably impact assembly plant capacity decisions, no doubt some jobs would be lost—including employment in supplier plants—and some workers would face transition costs. On the production side, it is notable that the U.S. and Mexico would be slightly harder hit than Canada in percentage terms, but in total units of production the sales decline south of the border would total 21,777 units. Given the compositional effects discussed earlier, it is no surprise to find that although all three importing regions benefit, the import gains go disproportionately to the E.U., which sees its imports increase by almost 25% versus only 7.7% for Korea.

2.5   References

Anderson, S.P. and A. De Palma (1992), "Multiproduct Firms: A Nested Logit Approach," *Journal of Industrial Economics*, Vol. 40, pp. 261-76.

Brambilla, Irene (2005), "A Customs Union With Multinational Firms: The Automobile Market in Argentina and Brazil," NBER Working Paper No. 11745, November.

Bresnahan, Timothy F. (1981), "Departures from Marginal-Cost Pricing in the American Automobile Industry," *Journal of Econometrics*, Vol. 17, pp. 201-27.

Bresnahan, Timothy F. (1987), "Competition and collusion in the American automobile industry: the 1955 price war," *Journal of Industrial Economics*, Vol. 35, No. 4, pp. 457-82.

Berry, Steven T. (1994), "Estimating Discrete Choice Models of Product Differentiation," *RAND Journal of Economics*, Vol. 25, pp. 242-62.

Berry, Steven T., James Levinsohn, and Ariel Pakes (1995), "Automobile Prices in Market Equilibrium." *Econometrica*, Vol. 63 (July), p. 841–90.

Berry, Steven T., James Levinsohn, and Ariel Pakes (1999), "Voluntary Export Restraints on Automobiles: Evaluating a Strategic Trade Policy." *American Economic Review*, Vol. 89 (June), p. 189–211.

Berry, Steven T., James Levinsohn, and Ariel Pakes (2004), "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, Vol. 112, No. 1, p. 68-105.

Brenkers, Randy and Frank Verboven (2006), "Liberalizing a Distribution System: the European Car Market," *Journal of the European Economic Association*, March.

Cardell, N.S., (1997), "Variance Components Structures for the Extreme-value and Logistic Distributions with Applications to Models of Heterogeneity," *Econometric Theory 13(2)*, 185-213.

De Melo, J. and P. A. Messerlin (1988), "Price, Quality and Welfare Effects of European VERs on Japanese Autos," *European Economic Review*, Vol. 32, pp. 488-511.

Esteban, Susanna and Matthew Shum (2005), "Durable Goods Oligopoly with Secondary Markets: the Case of Automobiles," Working Paper, August.

Feenstra, R. C. and James A. Levinsohn (1995), "Estimating Markups and Market Conduct with Multidimensional Product Attributes," *Review of Economic Studies*, Vol. 62, pp. 19-52.

Fershtman, Chaim and Neil Gandal (1998), "The Effect of the Arab Boycott on Israel: The Automobile Market," *RAND Journal of Economics*, Vol. 29, No. 1, Spring, pp. 193-214.

Goldberg, Pinelopi K. (1995), "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry." *Econometrica* Vol. 63 (July), p. 891–951.

Goldberg, Pinelopi K. (1994), "Trade Policies in the U.S. Automobile Industry," *Japan and the World Economy*, May, pp. 175-208.

Goldberg, Pinelopi K. and Michael M. Knetter (1999), "Measuring the Intensity of Competition in Export Markets," *Journal of International Economics*, Vol. 47, No. 1, February, pp. 27-60.

Goldberg, Pinelopi K. and Frank Verboven (2001), "The Evolution of Price Dispersion in the European Car Market," *Review of Economic Studies*, Vol. 68, No. 4, pp. 811-48.

Irwin, Douglas A. and Nina Pavcnik (2004), "Airbus versus Boeing Revisited: International Competition in the Aircraft Market," *Journal of International Economics*, Vol. 64, pp. 223-245.

Petrin, Amil (2002), "Quantifying the Benefits of New Products: The Case of the Minivan," *Journal of Political Economy*, Vol. 110 (August), p. 705–29.

Verboven, Frank (1996a), "The Nested Logit Model and Representative Consumer Theory," *Economics Letters*, Vol. 50, pp. 57-63.

Verboven, Frank (1996b), "International Price Discrimination in the European Car Market," *RAND Journal of Economics*, Vol. 27, No. 2, Summer, pp. 240-68.

3. Impact on FDI in assembly plants

This section assesses the potential impact of eliminating MFN tariffs discussed in section two above for new vehicles on the location decisions of auto assembly in North America, particularly in Canada. What will be the impact of Canada's bilateral or regional FTAs, those already in existence and those under negotiation (such as Canada-Korea) on other trading partners' location decisions in Canada?

A tariff on final vehicle imports provides incentives for foreign firms to establish local production capacity to avoid the tariff, so-called tariff-jumping. While at the margin the effect certainly exists, current tariff levels are sufficiently low and the overcapacity in the industry sufficiently large that we do not expect much of an impact. The expected cost of the elimination of the tariff on final vehicle imports is the product of the following four factors:

(1) Probability that a foreign firm will decide to build a new assembly plant in North America in the near future.
(2) If such an investment would take place, the probability that the elimination of the Canadian tariff would stifle the project.
(3) If such an investment would have taken place, the probability that a site in Canada would have been chosen over one in the U.S. or Mexico.
(4) Net benefit of an assembly plant to Canada.

Each of these four factors will be discussed separately in subsections 3.1–3.4. We will argue that (1) few new capacity additions in North America can be expected in the next decade; (2) the impact of Canadian trade policy on such FDI decisions is likely to be minor; (3) the likelihood of any future investment in North America assembly capacity going to Canada is lower than in more central locations; (4) a significant fraction of the value to the Canadian economy will be "lost" to the firm making the investment in the form of a subsidy to attract the FDI in the first place.

Moreover, if a change in Canadian trade policy is matched by a similar tariff concession abroad, the effect would also work in reverse, as discussed in Section 3.5. Tariff-jumping FDI

abroad would stop, potentially increasing investment in new or existing North American plants.

3.1 New capacity additions in North America

The probability that a firm will expand assembly capacity in North America beyond the currently announced expansion plans is fairly small. Table 3.1 indicates the number of assembly plants in operation over the last thirty years. Even though the production level in North America was higher in 2004 than in 1985 (see Figure 1.1A), the number of assembly plants has remained more or less constant. Canada and the North-East of the United States have seen a loss in plants, and more closures have been announced. Mexico and the U.S. South-West, on the other hand, have seen more plants open than close over the last decade; these areas have been particularly popular with transplants—foreign producers.

Table 3.1: North American assembly plants (1975-2004)

	1975	1985	1995	2004	announced
Total plants	68	85	88	84	
By country					
Canada	10	12	14	10	+1, -1
USA North/East	35	46	43	41	-3
USA South/West	18	18	18	22	+2, -1
Mexico	5	9	13	11	1
By ownership					
American1	66	79	70	65	-5
Asian	1	4	14	16	4
European	1	2	4	3	

Notes: 1 Includes plants now owned by DaimlerChrysler; Ford will announce assembly plant closures in January 2006, the expectation is 3-4
Source: Ward's Automotive Yearbook (various years) and Ward's Infobank (2004)

A net decrease in capacity in the coming years is expected as GM and Ford are likely to close more plants than the European or Asian producers will open. In terms of FDI for Canada, it does matter where the transplants will put their new plants. For foreign producers that operate only a few plants on the con-

tinent, it is often advantageous to locate them close by one an-other so they can share suppliers more easily. Moreover, now that most cars and light trucks are produced in smaller model runs in a single plant and shipped across the continent, econo-mizing shipping costs makes the centre of the continent relatively more attractive than Canada[32].

Table 3.2 lists each foreign firm currently selling vehicles in North America with their production and sales statistics for 2002. Firms are ordered by total sales; it is also indicated what fraction is satisfied by domestic production, and how many vehicles are imported.

Toyota has just completed construction of a compact pickup plant in Baja California, Mexico and will start production at its full-size pickup plant in Texas in 2006. It has announced a new plant for compact SUVs in Woodstock, Ontario. Production of each of these plants is not factored into Table 3.2 yet and they will add at least 400,000 vehicles to Toyota's North American production capacity. Given the high growth rate of Toyota's North American sales—it consistently averages almost 10% in Canada and the U.S. and it only recently entered the Mexican market—a new plant is certainly on the horizon.

Honda has also seen large sales increases, but in 2002 it only imported 334,000 vehicles. While this is certainly enough to fill an assembly plant, this total comprises a wide range of models that even with Honda's flexibility would be hard to produce in one plant. Honda now produces more vehicles in North America

---

[32] The changing geography of the industry in North America is a topic I cannot possibly do justice here. I refer the interested reader to recent work by Thomas Klier, senior economist at the Federal Reserve Bank in Chicago. His most recent analysis on the subject have appeared in the bank's *Economic Perspectives* series, third quarter of 2005. *Chicago Fed Letters* in February 2005 and March 2006 have featured articles on the transition of the auto supplier industry with a particular focus on the role of the Midwest. On the other hand, the analysis in the presentation that Sean McAlinden of the Center for Automotive Research gave at the April 2006 conference on "The New Geography of Auto Production" organized by the Chicago Fed in Detroit was much more critical of the North-South shift (the presentation is online at the web site of the Chicago Fed).

than in Japan. Its North American sales have increased a lot in the last three years as it entered many new segments, especially in light trucks. In the past, Honda added a second assembly line to an existing site (Marysville, OH and Alliston, Ontario) to increase capacity. Especially its latest plant in Lincoln and its Mexican plant are still a lot smaller than their other operations and could be expanded before Honda ventures to a new site.

Table 3.2: North American production and sales of foreign firms				
	Production		Sales	
		total	domestic	imported
Toyota[1]	1,196,019	1,912,729	1,110,753	801,976
Honda	1,138,717	1,443,595	1,109,618	333,977
Nissan[2]	750,925	1,016,167	714,512	301,655
Volkswagen	332,876	662,585	297,211	365,374
Hyundai		442,036		442,036
Mitsubishi[1]	174,466	349,200	170,268	178,932
Mazda1	47,603	329,353	120,151	209,202
BMW	124,374	280,295	58,662	221,633
Kia		266,359		266,359
Subaru-Isuzu[1]	131,833	255,438	137,912	117,526
Mercedes-Benz	102,983	231,315	43,337	187,978
Suzuki1	12,609	79,413	8,380	71,033
Daewoo		38,254		38,254
Porsche		22,793		22,793
Renault		15,386	11,185	4,201
Peugeot		9,148		9,148

Notes:
1 Production includes the shares in joint ventures: NUMMI (Toyota), AutoAlliance (Mazda), CAMI (Suzuki), Subaru-Isuzu, Diamond-Star (Mitsubishi). Includes Toyota production at NUMMI;
2 Nissan production includes its output for Renault in Mexico

The third largest transplant producer, Nissan, operates a huge plant in Smyrna, TN and two large Mexican plants. In 2002 it imported 302,000 vehicles and it is growing strongly recently. Given the closer integration with Renault, there are always rumours that the French automaker might consider a comeback to the U.S., but that is highly speculative.

After a number of lean years, Mazda is working its Auto-Alliance joint venture with Ford flat out, producing more than

260,000 vehicles at full capacity. The majority of these, however, are Ford Mustangs. Even with a string of very well received models, in 2004 the Mazda 3 became Canada's best selling passenger car, its total North America sales have not increased beyond its 2002 level. Given that Ford, which owns a controlling stake in Mazda has a lot of spare assembly capacity, greenfield investments by Mazda are not on the horizon.

Mitsubishi, Subaru, Isuzu, and Suzuki are in not in great shape and surviving is the first priority for these firms now. The proliferation of vehicles, discussed in Section 1.3, greatly increases the development burden for these smaller firms. Mitsubishi was associated with DaimlerChrysler and they were developing a compact car together, but that link has been severed. Subaru was partly owned by GM, but that stake was taken over by Toyota. It will take a long time to integrate production of Subaru's in Toyota plants, should Toyota choose to do so. Isuzu and Suzuki are still partly owned by GM, but their total sales would barely dent the surplus capacity at GM.

Summing up for the Japanese producers, Toyota is likely to increase its North American assembly capacity by at least one plant in the next decade. Honda and Nissan might consider investing as well, but their plant has to be either relatively small or extremely flexible, because their imports are a varied bunch. The recent decrease in the yen, the possibility to expand existing North American factories (outside Canada), and Honda's overcapacity in Japan, makes a new plant unlikely to happen soon.

The next investor in North American assembly plants will be Hyundai, which recently opened a plant in Alabama. Early 2006 it decided on a site in Georgia for its Kia subsidiary, nearby its Hyundai plant in Montgomery, AL so it can share suppliers for its two plants. Further capacity expansions are highly uncertain; the viability of the Kia plant already relies on a very ambitious sales projection and the Alabama plant will take some time to ramp up its production to its full capacity of 300,000 vehicles per year.

Finally, in 2003-04 the European producers were also considering North American assembly plants when the euro was breaking records on the currency markets on a daily basis. More recently, North American production capacity is not the highest

priority for most manufacturers. Volkswagen, the largest European importer, has seen sales of its main brand slump and faces high restructuring costs and overcapacity in Europe. An Audi plant is not entirely impossible, but at sales below 100,000 it is unlikely. Mercedes-Benz does not produce any sedans in North America and quality control problems makes this an unlikely proposition for the near future. BMW produces less than 10,000 cars in Mexico and around 25,000 roadsters in Spartanburg. The majority of its sedans and even its new compact SUV are imported. Given that it is unlikely that the Mini, the 3 series and its larger cars can be produced efficiently together in one plant, a new BMW plant in North America is also highly unlikely.

In sum, for the coming 10 years that leaves one plant for Toyota, probably one plant for Honda, and maybe one for Nissan to substitute domestic production for imports[33].

3.2 Sensitivity of investment in vehicle assembly to Canadian tariffs

If an assembly plant satisfies domestic content requirements under NAFTA, it qualifies for duty-free exports to other NAFTA member states. In this case, the effect of the elimination of the Canadian import tariff on a firm's likelihood to go ahead with the project is independent of the actual location of the plant—discussed in the next Section[34].

If a firm were contemplating constructing a new plant anywhere in North America (or add capacity to an existing plant), we have to consider how the probability of an investment

---

[33] At the time this document was last revised, both Honda and Kia had confirmed that they will build their next assembly plant in the U.S. Toyota's announcement to built a second assembly plant in Ontario almost guarantees no further assembly investments of the company in Canada in the near future. That leaves only Nissan as a possible new investor and according to news reports, the Ontario government has already started talks with the company (The Canadian Press, May 17, 2006).

[34] If a plant is located in Canada or Mexico, satisfying the domestic content requirements is imperative for a plant's viability because of the size of the U.S. market.

would be affected by a Canadian tariff cut. Given that production of virtually all North American plants is sold over the entire continent, Canadian sales will be proportional to Canada's size in the North American market. In 2002 this share was 8.7% and declining over time. Value weighted its share will be even lower as the average price of vehicles sold in the much larger U.S. market is higher and the lower average price in Mexico applies to only approximately half as many vehicles as sold in Canada.

For trade policy, the scenario we are interested in is a foreign producer that in the presence of the Canadian tariff of 6.1% would decide to build a North American plant, but in absence of the tariff would cancel the investment. For a typical greenfield investment of 150,000 vehicles that would mean on average 13,050 vehicles heading towards Canada. At a ballpark out-the-factory-door cost of $20,000, the elimination of the Canadian tariff would tilt the balance of costs and benefits of the new plant by less than $16m against investing in North America.

One can think of a lot of other idiosyncratic changes that would have an equal, even larger, effect. Note that the annual output of the hypothetical plant is estimated to be worth $3b and that each 1% change in the value of the foreign currency would have more than double the effect of the Canadian trade policy. To put this in perspective, over the last year the dollar has appreciated 15% against the yen and 12.5% against the euro. These trends are approximately 30 times as important as any change in Canadian trade policy. Any increase in shipping costs per vehicle by $105, which is likely to be less than the impact of the recent doubling in fuel prices, would have an equal effect in favour of locating in North America. A change in labour costs at the assembly plant of only 3.5% would also have a comparable effect.

In sum, we believe that the share of any North American production heading for the Canadian market is too small for the Canadian tariff of 6.1% to have much of an impact. Throughout, we have assumed that the U.S. tariff levels, at 2.5% for cars and 4% for light trucks, remain constant. Given the much larger importance of the U.S. market, the elimination of U.S. tariffs on final vehicle imports would have an impact on foreign firms' location decisions that is more than five times larger.

275

3.3 New capacity additions: Canada versus the U.S. or Mexico

Even if a firm would change its investment plans in response to a Canadian tariff cut, it would only constitute a loss of FDI for Canada if the plant would have been constructed there otherwise. As long as firms produce vehicles that satisfy the NAFTA domestic content requirements, currently 62.5%, the location within the NAFTA area is independent of the individual countries' tariff levels. Even though import tariffs on final vehicles in the U.S. are lower than in Canada and much lower than in Mexico, this does not make the U.S. a more attractive location because all local production can be traded within the area duty free.

Assuming a firm wants to establish a new assembly plant in North America (analyzed in Section 3.1) and assuming that this investment would be cancelled if Canada eliminated its import tariff (analyzed in Section 3.2), we now investigate what the probability is that Canada would have been chosen for a new assembly plant site. The Canadian track record in attracting FDI is readily available. Table 3.3 lists the ten most recently constructed or announced light vehicle assembly plants in North America. Only one of those plants will be built in Canada. The Northern U.S., the traditional hotbed of the industry only received two plants (one of which was very small). Clearly, the most popular region has been the Southern U.S. and to a lesser extent Mexico.

Three factors are important for future North American plants. There are clear network effects in organizing one's supply chain. If two assembly plants are located reasonably close, they can use the same supplier even for parts which are just-in-sequence, i.e., for which suppliers cannot be farther than a 2–4 hours drive. This makes it likely that the new Kia plant, which is supposed to be the next North American plant, will be constructed close to the Hyundai plant in Montgomery, AL. Such co-location decisions by foreign producers make it also more likely that their preferred suppliers from their home country will join them in North America.

276

Table 3.3: Location of the most recent light vehicle assembly plants in North America

Plant name	Owner	Start-up	Product
Canada			
Woodstock, ON	Toyota	2007	Compact SUVs
Northern United States			
Mishawaka, IA	GM (AM Gen.)	2001	SUVs
Lansing Gr. Rapids, MI	GM	2001	Cars and light trucks
Southern United States			
Lincoln, AL	Honda	2001	Light trucks
Canton, MS	Nissan	2003	Full size pickups
Montgomery, AL	Hyundai	2005	Cars and SUVs
San Antonio, TX	Toyota	2006	Full size pickups
Mexico			
Toluca	BMW	1999	Cars
Toluca North	DaimlerChrysler	2001	Light trucks
Baja California	Toyota	2005	Compact pickups

Source: Ward's Automotive Yearbook (various years); Automotive News (various issues)

Of course, given that both Toyota and Honda already own plants in Canada, this could work to Canada's advantage. In the case of Toyota, the ability to share suppliers with its well-established Cambridge operation was crucial in the selection of Woodstock, Ontario for its seventh North American plant. It is not implausible that Honda will also look at Ontario sites should it decide to build a new plant in North America. Furthermore, given that a lot of FDI takes the form of expanding an existing facility, the presence of two Toyota, one Honda, and one Suzuki-GM plant in Ontario also opens the door to further capacity increases in Canada. Table 3.4 contains all active, announced, and recently closed Canadian assembly plants.

Table 3.4: Canadian final assembly plants (light vehicles)

Plant name	Owner	Capacity	Start-up	Closed
Active:				
Alliston 1	Honda	390,000	1986	
Alliston 2	Honda		1998	
Cambridge North	Toyota	270,000	1988	
Cambridge South	Toyota		1998	
Ingersoll	CAMI (GM-Suzuki)	100,000	1989	
Oshawa Truck	GM	275,000	1964 (?)	
Oshawa #1	GM	545,000	1954 (?)	-1/3 in 2006
Oshawa #2	GM		1954 (?)	2007
Oakville	Ford	290,000	1953	
St. Thomas	Ford	230,000	1967	
Bramalea, (Brampton)	DaimlerChrysler	240,000	1986	
Windsor	DaimlerChrysler	350,000	1928	
Announced:				
Woodstock	Toyota	100,000	2008	
Closed:				
Ste. Therese, QU	GM		1965	2002
Ontario Truck (Oakville)	Ford		1965	2004
Pilette Road (Windsor)	DaimlerChrysler		1975	2003
Halifax, NS	Volvo		1963	1998
Bromont, QU	Hyundai		1989	1993

Notes: Unless otherwise indicated, all plants are in Ontario, this comprises all remaining plants.

A second factor, already mentioned before, is the desire to minimize shipping costs for vehicles, which tend to be much greater than for parts. As long as the most popular vehicles had annual sales greater than the minimum efficient scale of a single assembly plants, several plants around the continent were set up to satisfy demand. Currently, this is only the case anymore for a few full-size pickup trucks. Most other vehicles are assembled in a single North American plant. This makes a central location on the continent more attractive, and works to Canada's disadvantage.

The third factor that plays a large role in the selection of an assembly site is government subsidies. There is a large literature

on location incentive tournaments that pit multinational enterprises against the governments in whose jurisdictions they consider investing. It is uncertain to what extent the subsidies influence the investment decision, but they certainly have an impact on the location if the FDI goes ahead. The decision of Ford to completely overhaul its minivan plant in Oakville, which has been operating below capacity for a while, and Toyota's decision to locate its latest plant in Ontario were facilitated by the Ontario and federal governments' recent subsidy initiatives. In April of 2004, the provincial government made $500m in funds available under the Ontario Automotive Investment Strategy, to cover 10% of investment costs of projects exceeding $300m. The federal government launched the Canadian Skills and Innovations Project in June, 2004 and pledged $1b for Canadian manufacturing, half to match the Ontario initiative. While this indicates that the different Canadian jurisdictions are willing to enter the subsidy game to attract investment—making investments in Canada more likely—it also lowers the (remaining) value of an assembly plant to the economy, as the investing firm is able to extract some of the surplus.

Finally, it should be mentioned that a plant located in the U.S. could choose not to satisfy the NAFTA domestic content requirements and simply pay import duties on the (small) fraction of production that is exported to Canada. For example, the BMW plant in Spartanburg, NC and the Mercedes-Benz plant in Vance, AL are estimated to have only 35% domestic content, well below the 62.5% required for duty-free access to Canada. Given that the models built in these plants are less appealing to the Canadian or Mexican market, luxury SUVs and a roadster, the companies simply pay the import duties. In this case, the elimination of the Canadian tariff would not lead to lost FDI for Canada, but it would cost Canada tariff revenues.

3.4 Net benefit of a new vehicle assembly plant to the Canadian economy

Finally, in the unlikely case that a firm decides to cancel a Canadian investment project because the Canadian tariff was

cut, we discuss the loss to the Canadian economy of this lost FDI. This is a hotly debated topic and estimates range widely. The most recent paper on the topic, by Michael Greenstone and Enrico Moretti, estimates the spillover effect of such an investment on the regional economy from increases in local property values in the selected location relative to the trend in extremely similar runner-up location(s). They find a significant and positive effect using a sample of investment projects in a variety of manufacturing sectors.

The automobile industry is in some respects different from most industries. (i) With certainty additional employment will be generated in supplier plants that locate nearby, although the multiplier has been declining over time. (ii) Wage rates in the industry are substantially above manufacturing wages in similar locations and the difference seems too large to be explained entirely by human capital differences. The traditional explanation is that the well-organized unions have been able to extract some of the rents in this oligopolistic industry. (iii) The automobile industry is becoming increasingly high-tech. The R&D expenditure per capita in Michigan is the highest of any state in the United States and 85% of it is in automotive technology. Unlike many other sectors, the vast majority of research is privately funded. This research intensity can create technology spillovers to nearby firms and human capital spillovers in the workforce as workers receive continuous training.

While these factors would increase the beneficial effect for the local economy of attracting automotive investments, they have also encouraged governments to offer subsidies to attract these plants in the first place. While the size of the externalities associated with automotive FDI can be debated, they are certainly positive. However, the spillover effects would be positive for several jurisdictions. As a result, the competing jurisdictions will engage in a bidding war to attract the plant[35]. The winning jurisdiction does not have to give away the entire surplus, just enough to make the firm indifferent between itself and the next

---

[35] Maureen Molot-Appel (2005) discusses the subsidy games of the last two FDI waves in the automobile industry.

best alternative. The losing jurisdiction, however, should have offered the entire surplus it expected from the investment. As a result, the net gain to the Canadian economy of a successfully attracted FDI program is expected to be equal to the intrinsic value a Canadian location can bring to the firm.

Figure 2.1: An illustration of the optimal subsidy offer
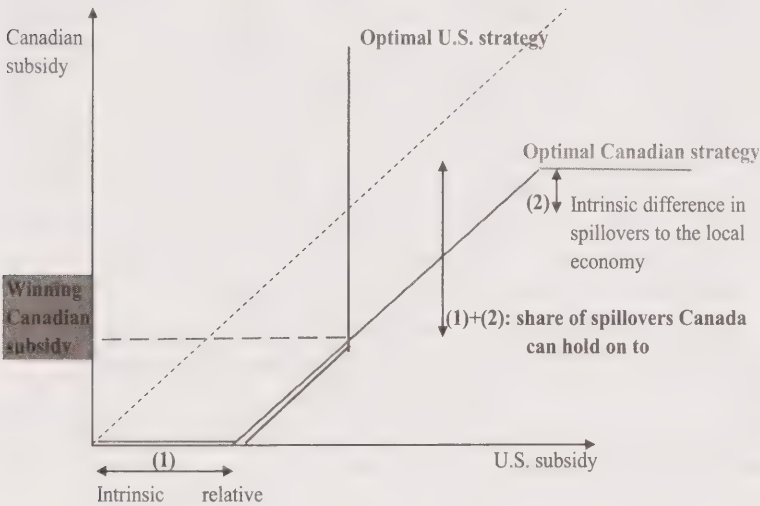
Figure 2.1 gives a graphical example of the equilibrium of such a subsidy game. In the hypothetical example depicted, a Canadian and U.S. jurisdiction compete to attract a new assembly plant by offering subsidies. Two magnitudes are important. (1) Indicates the intrinsic relative advantage of the Canadian site. As depicted, absent subsidies the firm would choose to locate in Canada. As long as the U.S. jurisdiction offers subsidies smaller than (1), where the blue line slopes upwards, Canada would be the preferred location. (2) Indicates the relative difference in spillovers to the local economy. In the example, this quantity would be greater for Canada, perhaps because of greater unemployment in the selected location, or more potential sites for suppliers to locate nearby. The maximum subsidy the Canadian jurisdiction will offer is equal to the total spillover it expects, indicated by the horizontal red line, and similarly the

maximum U.S. subsidy are its expected spillovers, the vertical blue line. Given that the blue U.S. line intersects the 45 degree line below the red Canadian line, we know that expected spillovers are larger for Canada, at least in this example.

The optimal subsidy offers are straightforward to derive. They are similar to the Bertrand Nash optimal price strategies discussed in most industrial organization books. The intersection gives the winning subsidy for Canada, $1 above the expected U.S. spillovers. Relative to the expected Canadian spillover, the winning jurisdiction is able to hold on to the magnitudes (1) and (2), its relative advantage for the firm and for itself over the next best alternative. Note that if the expected spillover would be larger in the U.S. jurisdiction, the relative size of (1) and (2) would determine the plant location and the winning jurisdiction would have to offer most of its advantage to the firm as a subsidy in order to attract the FDI.

The crucial insight to take away from this example is that even though the value to the local economy of automotive FDI might be very large, a significant fraction will accrue to the firm making the investment in the form of a subsidy to fend off competition from other jurisdictions.

3.5   Higher investment in Canada

Just as there is an ever so slight loss in FDI (in expectation) from the elimination of the Canadian imports, the same analysis can be applied in reverse if trading partners eliminate their tariffs. The expected benefit in this instance is the product of the same four factors considered above. Two differences are especially notable: foreign tariffs tend to be much larger, leading to a larger expected effect, but comparative advantage might disadvantage Canada in the relatively labour-intensive assembly stage of production, lowering the expected effect.

Exports of finished vehicles from North American are limited. Statistics in Table 3.5 are an attempt to construct export volumes for the major North American producers. While export statistics are not collected directly, we can obtain an estimate by subtracting sales of domestically produced vehicles from produc-

tion. The result is the sum of exports and inventory accumulation. Only one firm, Ford, is a significant exporter. This pattern is unlikely to change in the future; no one expects North America to become an export base for finished vehicles.

Table 3.5: Exports of light vehicles from North America (2002)

	Passenger cars		Light trucks		Total light vehicles	
	Production	Exports[1]	Production	Exports[1]	Production	Exports[1]
General Motors	2,458,052	54,284	3,159,053	104,438	5,617,105	158,722
Ford	1,437,905	442,105	2,690,958	366,641	4,128,863	808,746
Chrysler Group	649,673	20,983	2,042,153	130,826	2,691,826	151,809
Toyota	750,621	58,093	445,398	27,173	1,196,019	85,266
Honda	835,335	14,114	303,382	14,985	1,138,717	29,099
Nissan	544,026	40,997	206,899	-4,584	750,925	36,413
BMW	24,234	12,349	100,140	53,363	124,374	65,712
Mercedes-Benz			102,983	59,646	102,983	59,646
Volkswagen	332,876	35,665			332,876	35,665

1 Exports include changes in inventories from one year to the next.
Source: Own calculations based on Ward's Automotive Yearbook (2003)

The reverse effect of the preceding analysis would predict that lower tariffs in other countries would reduce the necessity for firms to set up plants overseas and instead satisfy foreign demand by exporting Canadian production. As in the previous section, it would be implausible to expect large effects. The modern flexible production systems are designed to sacrifice some scale economies, but increase the ability to produce a greater variety of vehicles on the same assembly line. This helps firms to offer a wider selection of vehicles in mature markets. At the same time, it also facilitates firms to produce vehicles closer to the final consumer. Rather than having each plant dedicated to a single vehicle, even assembly in smaller (overseas) markets becomes viable if a large fraction of the lineup that is sold in that country can be assembled in a single local plant.

The recent record confirms that the emerging demand for vehicles in Asia and Latin America is being met by adding local production capacity. The European market is increasingly being served from plants in Eastern Europe, where wages are substantially lower. Earlier statistics also illustrated the growing

importance of Mexico as a North American producer. While exports of finished vehicles from low-wage countries are only a marginal phenomenon for the moment, this might reverse in the near future. The labour cost in final assembly is too small to make up for the high transport cost of finished vehicles. Once the emerging producers develop their own supply chain and are able to produce a greater fraction of a vehicle's content domestically at lower wages, vehicle exports might take off. Hence, large export volumes of vehicles from Canada to the rest of the world seems an unlikely proposition.

4.   Market analysis: aftermarket components

This section estimates the potential impact of eliminating MFN tariffs on aftermarket auto parts on the production, employment, consumption, and trade of aftermarket auto parts in Canada; documents and analyzes the industrial trends in the auto parts sector; and assesses Canada's position and advantages.

4.1   Aftermarket parts

To put in perspective what part of the automobile industry we will discuss in this section, Table 4.1 contains total employment statistics for all the different automotive sub-sectors.

In total, the industry is estimated to have employed 892,700 workers in 2004, an increase of 6.8% over the 2000 total. Total employment breaks down into 45.8% in the after-market sector, 26.8% in manufacturing (parts and final assembly combined), 18% in vehicle sales, and a final 9.3% in other automotive related sectors, such as road construction or rentals[36]. The breakdown has been relatively stable over the last five years, although manufacturing continues its decline, mirroring the trend in the aggregate economy. The service side of the industry has clearly becomes the most important employer—the sum of employment in aftermarket and sales stood at 63.8% of the automotive workforce. Given that these services are largely non-tradable, the importance of trade policy has declined over time. In terms of total sales or value added, however, the contribution of manufacturing will be much higher than its employment share.

---

[36] Note that these statistics draw on Statistics Canada Labour Force Surveys and the totals are markedly different (higher) from the totals obtained aggregating employment statistics from the Annual Survey of Manufacturing. The relative importance of the different sub-sectors, however, is relatively similar using either source.

Table 4.1 Total employment in different sub-sectors of the automobile industry

	NAICS	2000	2004	Change 2004-2001
Aftermarket				
Motor Vehicles, Parts & Accessory wholesale	4152	23.3	29.1	24.9%
Automotive Parts & Accessories stores	4413	35.1	31.3	-10.8%
Gasoline Service Stations	4471	81.3	79.4	-2.3%
Other General Merchandise Stores	4529	98.8	120.1	21.6%
Automotive Repair & Maintenance	8111	140.1	149.2	6.5%
Total		**378.6**	**409.1**	**8.1%**
% of Automotive Employment		**45.3%**	**45.8%**	
Vehicle Assembly & Parts Manufacturing				
Motor Vehicle Manufacturing	3361	81.5	80.9	-0.7%
Motor Vehicle Body & Trailer Manufacturing	3362	20.1	18.4	-8.5%
Motor Vehicle Parts & Accessories mfg	3363	133.8	140.2	4.8%
Total		**235.4**	**239.5**	**1.7%**
% of Automotive Employment		**28.2%**	**26.8%**	
Wholesale & Retail Sales of Vehicles				
Motor Vehicles, Wholesale	4151	13	13	0.0%
Automobile Dealers	4411	134.3	147.7	10.0%
Total		**147.3**	**160.7**	**9.1%**
% of Automotive Employment		**17.6%**	**18.0%**	
Other Automotive				
Highway, Street and Bridge Construction	2373	57.4	60.8	5.9%
Automotive Equipment Rental & Leasing	5321	17.4	22.6	29.9%
Total		**74.8**	**83.4**	**11.5%**
% of Automotive Employment		**8.9%**	**9.3%**	
Total Automotive Industry		**836.1**	**892.7**	**6.8%**

Source: DesRosiers Automotive Yearbook 2005, Based on Statistics Canada Labour Force Surveys

An important feature of the automotive parts industry is that original equipment manufacturers (OEMs), i.e., the vehicle assemblers, can import parts duty free if they operate an assembly plant in Canada. Hence, trade policy affects only a fraction of the "Motor Vehicle Parts and Accessories Manufacturing" sector. All parts imports that are used in the assembly of new vehicles (OE parts) in North America are exempt from tariffs and hence unaf-

fected by trade policy changes. As such, the import competition Canadian firms face on OE parts will not increase if current tariffs on parts are abolished. To the extent that other countries have extended similar duty exemptions for imported parts that enter locally produced vehicles, Canadian exporters are already exempt from duties on their OE parts exports as well. In addition, Canadian exports of parts are predominantly going to the U.S.; due to NAFTA, these are also not affected by tariffs. This is likely to be the case for aftermarket parts as well.

Only parts that are used in repairs, maintenance, or upgrades of existing vehicle (in the aftermarket sector) are subject to duties. Trade with the U.S. or Mexico—the vast majority of Canadian imports or exports—falls under NAFTA and does not incur duties. Imports from countries with most favoured nation status[37]—the vast majority of trade outside of NAFTA—is subject to tariffs ranging from 0 to a maximum of 8.5%.

In summary, current Canadian import tariffs outside NAFTA are set at 0—8.5% for aftermarket parts, 6.1% for finished cars and light trucks, and 0% for OE parts. Note that such discrimination between finished vehicles, and OE or aftermarket parts is common. For example, China has for a long time provided incentives for local assembly by setting the tariff rate on finished vehicles much higher than for parts (as high as 100%). On April 1, 2005 it overhauled its import regime, classifying complete kits used to assemble vehicles locally with minimal domestic value added as completed vehicles, incurring on average a 30% tariff rate. Parts used in local assembly, on the other hand, only incur 15% tariffs. In its agreement with the WTO, China committed to lowering these rates to 25% and 10.3%, respectively, by mid-2006[38].

Employees listed in the first sector of Table 4.1, "Aftermarket", are mostly engaged in sales, administrative, repair, and maintenance tasks. They are using components but not manufacturing them. Employees making parts will be listed in the NAICS

---

[37] With China's inclusion in the WTO in 2001, all Canada's significant trading partners in the automotive industry now enjoy MFN status.

[38] Automotive News, March 28, 2005 "China closes importer's tax loophole."

3363 industry, which employed 140,200 workers in 2004, an increase of 4.8% over 2000. However, the vast majority of those parts is destined to OEM. customers and is largely immune to any trade policy changes because competing imports are not subject to tariffs. However, on the export market it will depend on the tariff treatment of OE parts in overseas destinations. In South Korea, these parts are taxed at 8% and in China at 15%[39]. In order to get an idea of the relative size of aftermarket components in the parts manufacturing industry, we now discuss some trends and summary statistics for the automotive aftermarket.

Table 4.2 lists total retail sales of parts (all figures in billions of CDN$) from 1990 to 2003, the last year for which complete data are available, for the domestic Canadian industry.

Table 4.2 Total retail sales of automotive parts and accessories (at retail prices, billion CDN$)

	Total retail sales of parts	Do-It-Yourself (DIY)	Installed		DIY	installed	labour
			parts	labour	(% of retail sales)		
1990	$10.52	$2.09	$4.64	$3.79	19.9%	80.1%	36.0%
1991	$10.68	$2.13	$4.66	$3.89	19.9%	80.1%	36.4%
1992	$11.40	$2.23	$4.96	$4.21	19.6%	80.4%	36.9%
1993	$11.66	$2.17	$5.09	$4.40	18.6%	81.4%	37.7%
1994	$12.16	$2.25	$5.31	$4.60	18.5%	81.5%	37.8%
1995	$12.27	$2.06	$5.47	$4.74	16.8%	83.2%	38.6%
1996	$12.49	$2.08	$5.59	$4.82	16.7%	83.3%	38.6%
1997	$13.17	$2.18	$5.90	$5.09	16.6%	83.4%	38.6%
1998	$13.08	$1.97	$5.94	$5.17	15.1%	84.9%	39.5%
1999	$13.38	$1.89	$6.09	$5.40	14.1%	85.9%	40.4%
2000	$13.89	$1.90	$6.31	$5.68	13.7%	86.3%	40.9%
2001	$14.40	$2.00	$6.49	$5.91	13.9%	86.1%	41.0%
2002	$15.07	$2.13	$6.79	$6.15	14.1%	85.9%	40.8%
2003	$15.45	$2.21	$6.93	$6.31	14.3%	85.7%	40.8%

Source: DesRosiers Automotive Yearbook (2005)

---

[39] A FTA with Korea or China is likely to boost Canadian parts exports. For the results in Table 4.13, we do not distinguish between aftermarket and OE parts because these countries add tariffs to either category.

Over thirteen years, the industry increased sales by 46.9% in nominal terms or 3% per year, cumulatively. This is more or less in line with the rate of inflation for car-related expenses, indicating quantity is almost entirely flat over this period. This is clearly not a growth industry. The increased number of vehicles on the road tends to raise sales, but improved quality of the existing fleet offsets this to a large extent. If customers hang on to their vehicles longer, given the increased reliability, the larger number of cars will eventually lead to greater aftermarket sales. At least in the short run, all net growth for the aftermarket sector has to come from exports.

Breaking down the total retail sales into the "do-it-yourself" (DIY) market, parts customers purchase from retail outlets, and components installed by professional repair stores or at dealerships reveals that the relative importance of the DIY market has declined noticeably. The total value of parts that customers purchased directly has hardly increased in thirteen years. Two-thirds of the increasing share of installed parts is the result of higher labour charges—see the last column of Table 4.2.

The sector is also becoming more competitive at the retail end. Limited to the last ten years of data (1993-2003), Table 4.3 lists aftermarket parts sales at retail (first column) and wholesale (second column) prices. While retail sales increased from $7.26 billion to $9.14, an increase of 25.9% or 2.3% per year, wholesale sales increased by 30.9% from $3.62b to $4.74b, 2.7% per year. The good news for Canadian parts manufacturers is that sales of parts grow more quickly than what one would assume based on retail trends. However, it also reveals that the total (off-the-factory) market is relatively small. In 2003 the total Canadian market for aftermarket parts at wholesale prices is only $4.74 billion dollars, less than a tenth of the total parts market (including OE parts). It also suggests that there is severe price pressure. Retail mark-ups declined, which can be gathered from the increasing ratio of retail to wholesale sales[40]. Finally,

---

[40] The importance of the installed parts sector leads to a much larger mark-up than we would expect if all sales were directly to customers—a ratio of 51.9% translates to a 93.7% mark-up.

the last two columns indicate that more than half of the increase in parts sales can be explained by a larger number of registered vehicles in Canada. Average parts sales per vehicle increased by only 1.1% per year in nominal terms, or $27 cumulatively over the ten year period. Note that parts per vehicle are an increasing function of the age of the vehicle. As a result, countries outside of North America, which tend to have much older fleets than Canada or the U.S., will have larger aftermarket sales per vehicle.

Table 4.3. Retail and wholesale aftermarket parts sales and sales per vehicle

	Total aftermarket parts sales (DIY and installed)		ratio	registered vehicles	Parts sales per vehicle
	(retail prices)	(wholesale prices)	(retail / wholesale)	(millions)	(wholesale prices)
1993	$7.26	$3.62	49.9%	15.509	$233
1994	$7.56	$3.77	49.9%	15.578	$242
1995	$7.53	$3.77	50.1%	15.871	$238
1996	$7.67	$3.87	50.5%	15.316	$253
1997	$8.08	$4.10	50.7%	16.076	$255
1998	$7.91	$4.03	50.9%	16.322	$247
1999	$7.98	$4.06	50.9%	17.071	$238
2000	$8.21	$4.20	51.2%	17.101	$246
2001	$8.49	$4.35	51.2%	17.668	$246
2002	$8.92	$4.60	51.6%	17.911	$257
2003	$9.14	$4.74	51.9%	18.207	$260

Source: DesRosiers Automotive Yearbook (2005)

Another important trend is the increasing importance of original equipment suppliers (OES), the dealerships associated with the vehicle manufacturers, see Table 4.4. The relentless price pressure in the final vehicle market has reduced dealers' profit margins on new car and truck sales. They have moved downstream and after-sales maintenance and repairs are making up a growing share of their sales. Vehicle companies only took 26% of the total aftermarket parts sales in 1993, but 35.2% in 2003. This has come predominantly at the expense of retailers, while warehouse and distributors—who mostly supply non-dealer installers of parts—have kept their market share at 37.3%. For Canadian suppliers, it means that firms like NAPA

or Uni-Select Canada are becoming relatively more important than Canadian Tire or Wal-Mart. Given that the former are industry insiders, while mass-market retailers are eventually marketing to customers directly, this changes the competitive position of Canadian firms relative to foreign competitors. It might be easier to bring factors other than price, such as quality or reliability of supply, into consideration.

Table 4.4. Aftermarket parts sales by distribution outlet

	Total aftermarket parts sales	warehouse /distributor	retail head offices	OES: vehicle companies	(as a % of the total market)		
	(wholesale prices)	(1)	(2)	(3)	(1)	(2)	(3)
1993	$3.62	$1.35	$1.32	$0.94	37.3%	36.5%	26.0%
1994	$3.77	$1.39	$1.37	$1.01	36.9%	36.3%	26.8%
1995	$3.77	$1.39	$1.33	$1.05	36.9%	35.3%	27.9%
1996	$3.87	$1.41	$1.35	$1.11	36.4%	34.9%	28.7%
1997	$4.10	$1.48	$1.31	$1.21	36.1%	32.0%	29.5%
1998	$4.03	$1.47	$1.26	$1.25	36.5%	31.3%	31.0%
1999	$4.06	$1.51	$1.31	$1.29	37.2%	32.3%	31.8%
2000	$4.20	$1.57	$1.26	$1.37	37.4%	30.0%	32.6%
2001	$4.35	$1.64	$1.25	$1.47	37.7%	28.7%	33.8%
2002	$4.60	$1.73	$1.29	$1.58	37.6%	28.0%	34.3%
2003	$4.74	$1.77	$1.30	$1.67	37.3%	27.4%	35.2%

Source: DesRosiers Automotive Yearbook (2005)

Finally, in Table 4.5 we compare the total size of the automotive aftermarket parts market with two benchmarks. In the first column, we replicate the aftermarket retail sales, studied in Table 4.2. The same market at wholesale prices, as in Table 4.3, is replicated in the second column. This is contrasted with the total market for OE parts—those sold directly to vehicle assemblers for installation in new vehicles. In Canada, this market has increased from $45.10 billion in 1998 to $48.09 in 2003. Comparable figures for the North American market, which is almost seven times larger, are in the fourth column. These numbers represent approximately 68% to 70% of the total value of shipments from automobile assembly plants. As Canadian parts suppliers are very much integrated into the North American

automobile industry, they compete in a $304.92 billion CDN market when they supply OEMs directly. On the other hand, if they focus on the aftermarket, the entire North American market is likely to be at most one fifth—estimated at $61.7 billion CDN based on U.S. and Mexican registrations[41].

Table 4.5. Relative size of different automotive parts sectors (billion CDN$)

	Demand				Supply	
	Aftermarket parts (Canada)		OE parts		NAICS 3362	
	(retail prices)	(wholesale prices)	Canada	North America	Canada	US
1998	$7.91	$4.03	$45.10	$315.41	$26.04	$273.60
1999	$7.98	$4.06	$55.42	$370.01	$28.58	$301.32
2000	$8.21	$4.20	$56.05	$359.69	$29.89	$303.48
2001	$8.49	$4.35	$49.70	$345.41	$28.59	$289.74
2002	$8.92	$4.60	$50.29	$350.74	$29.69	$317.80
2003	$9.14	$4.74	$48.09	$304.92		

Source: DesRosiers Automotive Yearbook (2005)

When we analyze the impact of trade policy, no separate data are available for OE and aftermarket parts. We only observe trade flows of total parts and also at the industry level we only observe manufacturing for all parts combined (NAICS 3362). Sales of the parts manufacturing sector are listed in the last two columns of 4.5. A couple of ratios should be remembered:

- Aftermarket parts sales in Canada make up 8.97% of the domestic demand for automotive parts (including OE parts).
- At the North American level the comparable fraction is much higher, most likely in the 15–20% range (as Canada has a disproportionate demand for OE parts).
- Canadian parts production only covers 54.1% (in 2002) of domestic demand for parts. In contrast the U.S. and Mexico both run parts surpluses.

---

[41] In absence of comparable data to those in column 2 of Table 4.5 for the U.S or Mexico, we estimate the total North American aftermarket by multiplying the average cost of automotive parts per vehicle for Canada, in Table 4.3, by the total registrations in North America.

4.2 Parts manufacturing (NAICS 3363)

When we study the impact of a changed trade policy, we will focus on NAICS industry 3363: "Motor Vehicle Parts & Accessories Manufacturing." Table 4.6 presents total employment statistics for the different sub-sectors.

Table 4.6. Total employment in motor vehicle parts and accessories manufacturing

| NAICS | | Total employment | | |
		1998	2002	change
3361	**Motor Vehicle Manufacturing**	51,440	47,495	-7.7%
3362	**Motor Vehicle Body & Trailer Manufacturing**	17,502	19,528	11.6%
3363	**Motor Vehicle Parts & Accessories Manufacturing**	94,264	88,840	-5.8%
336310	Motor Vehicle Gasoline Engine and Engine Parts Manufacturing	10,227	10,522	2.9%
336320	Motor Vehicle Electrical and Electronic Equipment Manufacturing	6,565	6,366	-3.0%
336330	Motor Vehicle Steering and Suspension Components (except Spring) Manufacturing	6,616	4,792	-27.6%
336340	Motor Vehicle Brake System Manufacturing	7,671	6,556	-14.5%
336350	Motor Vehicle Transmission and Power Train Parts Manufacturing	11,090	9,886	-10.9%
336360	Motor Vehicle Seating and Interior Trim Manufacturing	13,130	12,598	-4.1%
336370	Motor Vehicle Metal Stamping	16,133	13,255	-17.8%
336390	Other Motor Vehicle Parts Manufacturing	22,832	24,835	8.8%
3361–63	**Automobile industry (manufacturing)**	163,206	155,863	-4.5%

Source: DesRosiers Automotive Yearbook 2005, based on Statistics Canada Census of Manufacturers

Note that the aggregate statistics differ from those in Table 4.1. While the former are compiled from labour market surveys, the statistics in the current table are the sum of employment of all establishments assigned to the NAICS 3363 industry. Plants that produce more other goods than automotive parts (for example assembly plants where some parts are produced on site) are

excluded here, while workers could still denominate their industry as NAICS 3363.

The total parts manufacturing sector employs 87% more workers than vehicle assembly (NAICS 3361): 88,840 versus 47,495. The most important sub-sector, increasingly so over time, is "Other Motor Vehicle Parts Manufacturing", while stamping and seats and interior trim are also very important in Canada. The only two sub-sectors that increased employment from 1998 to 2002 are engines & engine parts and "other parts". By 2002, the latter accounts for 28% of employment in the sector.

Table 4.7. Summary statistics on the motor vehicle parts and accessories sector

NAICS		employment share		production workers in employment (%)		value added per production worker		value added per hour worked (prod.)	
		1998	2002	1998	2002	1998	2002	1998	2002*
3363	Parts	100%	100%	86.5%	85.3%	$139,232	$144,496	$65.34	$67.81
336310	engine	10.8%	11.8%	85.2%	88.5%	$337,637	$177,734	$143.06	$75.31
336320	electrical	7.0%	7.2%	88.1%	83.2%	$117,731	$101,952	$57.86	$50.11
336330	steering & suspension	7.0%	5.4%	87.2%	88.0%	$118,823	$150,977	$55.49	$70.50
336340	braking	8.1%	7.4%	87.8%	84.8%	$123,256	$117,703	$60.45	$57.73
336350	transmission	11.8%	11.1%	87.4%	83.4%	$143,750	$191,659	$62.83	$83.77
336360	interior	13.9%	14.2%	87.7%	87.2%	$120,589	$149,695	$58.77	$72.96
336370	stamping	17.1%	14.9%	86.2%	85.7%	$109,222	$141,341	$53.61	$69.38
336390	other	24.2%	28.0%	85.0%	83.6%	$98,447	$126,541	$46.42	$59.67

Source: own calculations based on DesRosiers Automotive Yearbook (2005) and Statistics Canada

In Table 4.7 we calculate a number of crucial summary statistics for the different sub-sectors. The share of production workers in employment is declining slightly over time in most sectors. This trend is particularly pronounced in electrical & electronic equipment manufacturing and the important "other parts" sub-sectors and could indicate an increasing level of technical sophistication. The average value added per production worker increased only slightly from $139,232 in 1998 to $144,496 in 2002, but this masks big differences between sectors. The highly capital intensive engine & engine parts sector

operated in 2002 at less than 53% of its value added of 1998. Inspection of the different components that enter these calculations reveals that this is the result of increased purchases of parts and materials, up 58%, while employment and total shipments increased only slightly.

Two more sectors saw their value added per worker slump, albeit not as drastically. The electrical & electronic equipment and braking sectors saw (nominal) declines of 13.4% and 4.5% respectively.

Two factors contributed positively towards aggregate value added growth for the industry. First, even though the engine sector saw its labour productivity decline, it is still above the industry average. Given that it increased its share of parts employment from 10.8% to 11.8% this relative reallocation of workers increased aggregate productivity. Second, labour productivity growth is positive and large in the remaining sectors. Average productivity growth is 28.5% over four years or 6.5% per year, if it is positive, ahead of the average growth for the manufacturing sector. The steering & suspension, transmission, and stamping sectors achieved this productivity growth by reducing total employment and keeping value added more or less constant. The seats & interior trim kept its employment constant, but increased its share of total parts manufacturing employment. Finally, the "other parts" industry increased labour productivity by 28.5%, while at the same time employing 1350 more workers. Finally, in the rightmost columns we calculate labour productivity per hour worked where the same trends are apparent. Normalizing by the average number of hours worked in each sub-sector, the same trends show up, but the dispersion across industries at each point in time is reduced.

4.3   Threats and opportunities

4.3.1   Threats

Before analyzing the impact of trade policy changes, we briefly discuss the most pressing threats and opportunities faced by the Canadian automotive parts industry. We draw from industry

coverage in the weekly *Automotive News* magazine and a survey of Canadian parts manufacturers organized jointly by the Canadian Auto Parts Manufacturers Association (APMA) and the Asia Pacific Foundation of Canada[42].

The most obvious threat to Canadian firms in the last couple of years is the appreciating Canadian dollar. The Canadian-U.S. **exchange rate** moved from a low of 0.618 on January 21, 2002 to a high of 0.885 on March 2, 2006. That is an appreciation of 43.2% in only four years. Given that the U.S. is by a wide margin Canada's largest trading partner this matters a lot. In addition, some export contracts to other countries are likely to be priced in USD as well.

While such exchange rate movements make Canadian suppliers less competitive internationally, they also hurt firms on their existing contract if **contractual prices** are specified in U.S. dollar terms. With volatile exchange rates, raw material prices (see below), and customers' market shares (see below), it becomes crucial not to be pinned down by fixed nominal prices in long term contracts.

Prices of raw materials, especially steel and oil, but other metals as well, have also increased enormously in recent years. These have been cited in most prominent U.S. Chapter 11 bankruptcy filings, e.g., Tower Automotive and Collins & Aikman. Of course, the extent to which firms are affected depends again on whether they are compensated for this. In the case of Tower Automotive, its customers have only agreed to change the agreed prices to reflect higher input costs of steel after the firm filed for bankruptcy.

Market share of traditional (Big 3) customers[43] is declining. Most domestic suppliers tend to have a disproportionate exposure to these firms. In addition, each of these firms is run-

---

[42] "The East Asian Automobile Industry: Opportunity or Threat? Results of a Survey of the Canadian Auto Parts Manufacturers," Canada in Asia report, January 2005.

[43] Throughout, we will be referring to the "Big 3" to indicate GM, Ford, and DaimlerChrysler; traditionally the largest three OEM customers in North America.

ning a very aggressive cost-cutting program to reduce their input purchase bills.

For example, in 2004 it took Ford almost a year to convince some of its global suppliers to sign the new contracts. In addition to severe price cutbacks, major concerns were raised by stipulations in the contract that allowed Ford to hand R&D work over to competitors. **Ownership of IP** is quickly becoming an important part of sourcing relationships. For example, in 2005 Multimatic sued Faurecia alleging it sourced a proprietary design with a competitor to get a lower price. With the rise of China's manufacturing capabilities, conflicts over IP are rising overseas as well.

A more general problem is that contracts often specify a fixed price per part. Firms spread out **fixed investment costs** over the expected model run. If a model proves less successful than anticipated, the contracted average price will not allow firms to recover their fixed setup costs. As suppliers are shouldering more of the R&D burden, this problem is becoming more widespread.

Not surprisingly, **import competition** was also cited by Canadian suppliers as one of the most significant threats. The two countries mentioned most often were the U.S. and China.

FDI by foreign suppliers who followed OEM transplants to North America is also perceived as a threat by domestic firms. For example, even before Toyota announced its intention to build a new assembly plant in Woodstock, Ontario, its seat supplier Araco of Japan opened a seat and interior trim plant in Ontario to keep up with Toyota's production expansion in North America. Similarly, Honda suppliers like Musashi Seimitsu (suspension and steering parts) or Ube (wheels) are expanding in Ontario to serve customers besides Honda as well.

On the export front, many firms also mention **difficulties exporting to the U.S.** This covers both concerns about insufficient (government) investments in border infrastructure as well as rising U.S. protectionism in the wake of 9/11.

4.3.2   Opportunities

Luckily, not everything is grim in the outlook of Canadian suppliers. One of the most frequently cited opportunity is the **expansion of production overseas**, especially in China. The vast majority of Canadian suppliers, 64%, have been asked by their North American clients to expand production capacity overseas to serve them better in locations abroad. Given that net growth in worldwide demand for new vehicle is predominantly outside of North America, it clearly makes sense to focus on developing countries.

In addition, many firms perceive foreign expansion as a way to attract new customers. The expectation is that these relationships will eventually translate into new supplier contracts to **North American transplants** as well.

In contrast to what one would gather from reports in the media, lowering costs is not one of the main motivations to **increase foreign sourcing**. Access to important customers, strategic geographic positioning, and strategic fit (patents, R&D, staff) all rank higher than cost considerations.

In contrast with the focus on **Asia**, which more than 70% of Canadian suppliers perceive as an opportunity, the Free Trade Area of the Americas does not seem high on the radar screen, neither as a threat nor as an opportunity.

As mentioned in section 1, Canada has considerable expertise in **fuel cell** research. In addition, from early on the R&D efforts have been well plugged in to the automotive industry. Once the technology is ready for prime time, Canada is expected to be one of the major players (in sharp contrast with the hybrid technology which will only come to Canada in 2010).

The Canadian industry has frequently argued that in order to be globally competitive and form a thriving industry it needs a solid base. Most directly, expansion of **assembly capacity** of final vehicles in Canada was a prime objective—communicated very clearly by the Canadian Automotive Partnership Council. The investments by Ford in Oakville, the commitment to keep St. Thomas open (at least in the short run), the current negotiations to avert some of the announced capacity reductions at

GM's Oshawa facility, and especially Toyota's announcement to build a new plant in Ontario are all very encouraging.

OEMs are **increasingly outsourcing** more steps in the production process. This can both be a threat and an opportunity for suppliers. As mentioned earlier, the success of these outsourcing relationships hinges crucially on the contractual details. For example, DaimlerChrysler is compensating Karmann when the demand for the Crossfire vehicle that Karmann assembles fell off more quickly than expected. Without such guarantees, outsourcing becomes a risky endeavour for suppliers. If structured properly, these outsourcing trends open up new growth opportunities in an otherwise mature industry.

4.4 Industry structure

4.4.1 Exit

The difficulties in the industry have led to a large number of bankruptcies in the U.S., where 138 of the top 150 North American suppliers have their headquarters (including regional headquarters). Table 4.8 lists the thirteen largest bankruptcies by U.S. suppliers in the past 5 years. Eight U.S. companies that were among the 100 largest OEM parts suppliers worldwide filed for Chapter 11 restructuring, which is almost a quarter of the 34 U.S. companies on the list. A further five Tier 1 suppliers went under that did not make the global list, but are listed (or at least have been recently) on the list of 150 largest North American suppliers. Note that these companies often suffered a couple of years of declining sales before filing, at least relative to more successful companies, and their 2004 rank understates their importance. For example, Amcast ranked as high as 82 in North America in 1993 and 122 in 2003, but fell off the list in 2004.

Among the list of bankrupt companies, firms headquartered in Michigan are very prominent, firms producing steel-intensive products, stampings, castings, or frames, are also overrepresented. Federal-Mogul and Hayes Lemmerz International filed a couple of months after the 9/11 terrorist attacks when the U.S. seemed heading for a recession and these companies had trouble

servicing their debt. The other bankruptcies are more recent, with filings accelerating at the end of 2004. Total output in the North American automobile industry was still going strong—it still is—but these firms were exceedingly exposed to raw material prices, the sales decline of the traditional Big 3 (U.S.) OEMs, and import competition.

Table 4.8. Recent major bankruptcies by U.S. automotive parts suppliers

Firm	Headquarters	filed for Chapter 11	World-wide rank (2004)	Global OEM parts sales	Employ-ment
Delphi Corp.	Troy, MI	Oct-05	1	$28.60	185,200
Dana Corp.	Toledo, OH	Mar-06	15	$9.06	46,000
Collins & Aikman Corp.	Troy, MI	May-05	20	$3.98	23,000
Federal-Mogul Corp.	Southfield, MI	Oct-01	39	$3.35	42,000
Tower Automotive Inc.	Novi, MI	Feb-05	45	$3.20	12,000
Hayes Lemmerz International	Northville, MI	Dec-01	64	$2.00	11,000
Oxford Automotive	Troy, MI	Dec-04	98	$1.08	3,800
Meridian Automotive Systems	Dearborn, MI	Apr-05	99	$1.03	5,900

Firm	Headquarters	filed for Chapter 11	NA rank (2004)	Global OEM parts sales	Employ-ment
J.L. French Automotive Castings Inc.	Sheboygan, WI	Feb-06	102	$0.50	1,800
Intermet Inc.	Troy, MI	Sep-04	68	$0.70	5,200
Citation Corp.	Troy, MI	Sep-04	79	$0.60	5,200
Key Plastics, LLC	Northville, MI	Mar-00	132	$0.70	4,000
Amcast Industrial Corporation	Fremont, IN	Dec-05	N/A	$0.42	2,600
Total				**$55.22**	**347,700**

Notes: Ranks are from the Automotive News list of 100 largest suppliers worldwide or 150 largest suppliers in North America. Sales are global OE parts in 2004 in billions of USD. Employment statistics are the latest available (generally 2005).

The scale of this wave of bankruptcies is unprecedented. The thirteen largest filings accounted for annual sales of $55.22 billion USD in 2004 and an even higher volume in the preced-

ing years. This number only includes OE parts sales, with total sales often much larger still. At the time of this writing, these companies employed a total of 347,700 employees[44]. While it is unlikely that much of the production capacity and associated jobs will be liquidated, many workers will have to transition into new jobs and have to take pay cuts.

Furthermore, this is certainly not the end of the restructuring in the parts industry. Lear Corp., the 5th largest supplier in North America, is struggling to make its debt payments and its market value plunged by more than 60% in the preceding year as more analysts see a Chapter 11 filing as a distinct possibility. Lear employs 115,000 workers worldwide. A bankruptcy by Visteon, currently employing 49,000 full time employees, was only narrowly avoided courtesy of a very generous payout package by its former parent. Some plants were transferred back to Ford, which is trying to sell them off, and some workers also transferred back to Ford. While the old Visteon employed 70.2% of its workers in the U.S., the restructured Visteon counts 56.1% Mexican workers in its hourly workforce. Only 12 of its plants remain organized by the UAW, the principal labor union in the U.S. automotive industry[45]. Delphi counts approximately 60,000 hourly employees at its 50 Mexican plants.

The bankruptcy of Delphi, the largest supplier worldwide, is likely to have wide-ranging effects on the industry. The tiered organization of the supply chain means that financial problems at a large Tier 1 supplier trickle down to the next levels very quickly. If Delphi defaults on some of its trade credit, some of its suppliers that are already stretched by the increased competition would have a hard time surviving. The wave of Tier 1 firm bankruptcies in Table 4.8 has led a lot of smaller suppliers to file for Chapter 11 restructuring as well. An incomplete list is in Table 4.9.

---

[44] Delphi employs 185,200 employees worldwide and 76,000 in the U.S. 42 of its U.S. corporate entities are involved in the restructuring.

[45] Automotive News, June 20, 2005, "New Visteon has Mexican flavor".

Table 4.9. Recent bankruptcies by smaller U.S. automotive parts suppliers

Year	Smaller suppliers filing for bankruptcy
2006	Hastings Manufacturing Company
2005	American Remanufacturers Inc.; Allied Holdings, Inc; Metalforming Technologies Inc.; Trim Trends Co. LLC.; BBi Group
2004	Andover Industries
2003	Liteglow Industries, Inc.
2002	Harvard Industries, Inc.
2001	Rankin Automotive Group, Inc.; Valeo Electrical Systems, Inc
2000	Cambridge Industries, Inc.; Safety Components International, Inc.; Dorsey Trailers, Inc.; Safelite Glass Corporation

Another notable outgrowth of the supplier distress is that Delphi managed to negotiate a two-tiered wage system in its 2003 labour contract negotiations with the UAW[46]. This allows the company to pay new workers lower wages than its existing employees. The union has always resisted such discrimination, even though the OEMs have repeatedly pushed for this as well. It remains to be seen how important this change will turn out to be in practice.

Quite remarkably, Canadian firms have survived this carnage almost scot-free. Given the increased cost pressure induced by the adverse exchange rate movement, the rising importance of the U.S. border in the post 9/11 world, and the traditional over-exposure to the Big-3 traditional customers for Canadian firms, one could have expected even more problems than in the United States.

4.4.2   Concentration

While a lot of large firms are in financial difficulty, total industry concentration has clearly been increasing. Figure 4-1 plots the evolution in the share of the total material cost for the motor ve-

---

[46] Automotive News, September 29, 2003, "UAW gives Delphi half a loaf."

hicle assembly industry (NAICS 3361) in the U.S. and Canada combined that is accounted for by different groups of firms[47].

Figure 4-1 Top suppliers' sales as a fraction of total material cost in NAICS 3361 (US + Canada)
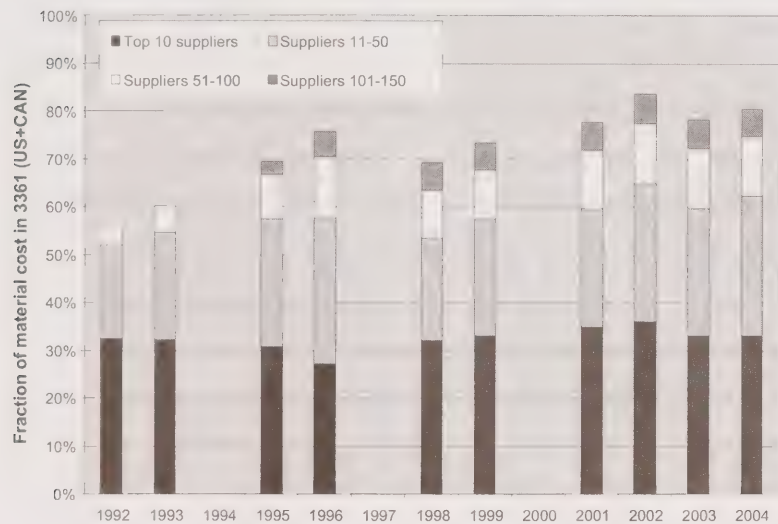
The share of OE parts sales in North America by the 100 largest suppliers, the longest time trend is available for this group, increased from 57% in 1992 to 75% in 2004. The bottom bars in each column, which track the sales of the top 10 suppliers, indicate that the increase is not caused by the very top firms. They kept their share approximately constant at one third. The largest increase is for firms in the second group, suppliers ranked 11 to 50 on the North American supplier list. This group increased its share from 19.4% to 29.2%, an increase of almost 50%. In order to break the top 50 in 2004, a firm had to sell $1 billion worth of OE parts.

In 1992 this would have secured the 21st spot on the list! Given the very moderate price increases in the industry as

---

[47] Given that the share of Mexico in North American production has increased over time, the increase in concentration for the total North American industry will be slightly lower, but the difference will be very small.

documented earlier—firms are clearly becoming much bigger. The next group of suppliers, ranked 51 to 100, only sold 4.8% of North American automotive materials, but 12.5% in 2004—an increase of 160%.

These patterns have two important consequences. In the tiered supplier system that is now in effect, only a couple of hundred firms have access to the final vehicle manufacturers. They decide which firms to outsource components to further down the line. For the vast majority of firms in the industry, relationships with these large tier 1 suppliers are crucial. Among the 150 largest suppliers were 9 Canadian firms in 1999, but only 7 remain. Four are in the top 100, versus 5 in 1999. The enormous expansion of Magna, which moved from the 7th spot in 1993 to 3rd in 2004, has increased the Canadian share on the list, but at the same time it concentrates a lot of Canadian automotive employment in a single company.

I have documented elsewhere[48] that the location of supplier headquarters is increasingly concentrated in Michigan. Several firms that used to operate regional headquarters in Canada have centralized their headquarters activities in their U.S. headquarters. Moreover, the number of Mexican, European, and Japanese suppliers on the list has crept up over time as well. All these factors imply that decisions about (overseas) outsourcing are increasingly being made outside of Canada, even though these decisions have an important impact on the Canadian industry.

An important reason for suppliers to become bigger has been the larger role in R&D that they have assumed in recent years. In order to diversify risk and to spread fixed costs of the development of new technologies over a larger volume of sales, scale is important. The relatively diminished role of Canadian firms in the upper echelons of the parts sector is likely to have an impact on the extent to which innovative activities are carried out in Canada.

---

[48] Van Biesebroeck J. (2006), "Trends and Complementarities in the Canadian Automobile Industry, (*forthcoming*) in Z. Chen and M. Duhamel, Industrial Economics and Performance in Canada.

4.5 Import demand and export supply

4.5.1 Elasticities

In order to assess the effect of trade liberalization on the domestic Canadian market and the export potential of Canadian firms overseas, we need estimates of the demand and supply elasticities for the different sub-sectors of the automotive parts industry. Methods to estimate demand elasticities in differentiated goods market developed in the industrial organization literature tend to be too data intensive to be widely applicable. In section 3 we estimated such a model for the Canadian final vehicle market and the data requirements clearly exceed what can conceivably be obtained for the parts industry.

In recent years, the international trade literature has also taken the fact that goods are differentiated more seriously and more reliable elasticity estimates are obtained exploiting properties of the constant elasticity of substitution (CES) demand system. A paper by Feenstra (1994) pioneered the approach and the estimation method was subsequently refined and applied to a much wider range of industries (products) by Broda and Weinstein (2006)[49]. The benefit of this approach is that demand estimates control explicitly for heterogeneity across goods, albeit in a restrictive way. The method can also deal with increasing variety and with quality or taste differences across goods or country of origin.

The estimates have been used to calculate the value of increased variety as an additional gain from trade. For example, Broda and Weinstein (2006) estimate the contribution of unmeasured growth in product variety in U.S. imports between 1976 and 2001 to be approximately 2.6% of GDP. They also find that the "true" import price index increases 1.2% per year more slowly—approximately one quarter of the annual increase—because of the increase in variety.

---

[49] Feenstra, R. (1994), "New Product Varieties and the Measurement of International Prices," *American Economic Review*, 87 (1), March, pp. 157-177 and Broda, C. and D. E. Weinstein (2006), "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 121(2), *forthcoming*, May.

A similar exercise on the export side, see Feenstra and Kee (2004, 2005), models a nation's output using a CES cost function that is decreasing in the number of varieties. An increase in variety in a sector will raise the sectoral price index and draw resources to the industry. An empirical application finds that in a cross section of countries, productivity levels are positively correlated to the number of varieties that are exported to the U.S. Over time, the relative evolution of a nation's productivity level is found to be similar to the evolution of its variety in exports.

For details on the methodology, we refer the interested reader to the papers by Feenstra (1994) and Broda and Weinstein (2006). Here we just provide a brief explanation of the underlying theory and some details on our implementation. Underlying the theory is a three-tiered CES utility function. At the upper level, consumers have preferences over two composite goods, one domestically produced and one imported:

$$
U_t = \left(D_t^{\frac{\kappa-1}{\kappa}} + M_t^{\frac{\kappa-1}{\kappa}} \right)^{\frac{\kappa}{\kappa-1}}, \quad \kappa > 1.
$$

D is the domestic good and the import composite, M, will be defined below. κ is the elasticity of substitution between the two goods. If this is equal to the elasticity of substitution between different imported varieties, the upper nest disappears[50].

In the second tier, the composite import good is defined as:

$$
M_t = \left(\sum_{g \in G} M_{gt}^{\frac{\gamma-1}{\gamma}} \right)^{\frac{\gamma}{\gamma-1}}, \quad \gamma > 1
$$

where M_{gt} is the sub-utility derived from the consumption of the imported good g in time t; γ denotes the elasticity of substitution among the imported goods, and G is the set of all imported goods.

---

[50] This is what one has to assume to use the model to investigate the impact of trade liberalization on the domestic industry; otherwise consumers allocate a fixed proportion of their budget to domestic and imported goods.

At the most detailed level, different varieties (c) are imported of each good (g) and we can use the non-symmetric CES utility function to define M_{gt}:

$$M_{gt} = \left(\sum_{c \in C} d_{gct}^{\frac{1}{\sigma_g}} m_{gct}^{\frac{\sigma_g - 1}{\sigma_g}} \right)^{\frac{\sigma_g}{\sigma_g - 1}}, \quad \sigma_g > 1 \quad \forall\, g \in G$$

For each good, imports are treated as differentiated across countries of supply. Tastes for varieties can differ or, alternatively, d_{gct} can represent heterogeneous quality levels for imports coming from different countries.

A major attraction of this framework is that one can construct a price index using the theory of exact index numbers without having to estimate the different taste or quality parameters (d_{gct}), only the elasticities of substitution. If a good is consumed (imported) even though its quality or desirability is lower than that of other goods, this difference has to be reflected in the price. Using expenditure shares one can aggregate prices to construct an exact price index. Note that for the demand system to be well-behaved, all elasticities of substitution have to exceed unity.

The demand system can be manipulated to find an explicit expression for the import demand equation for each good. Differentiating with respect to time gives the following demand equation in first differences:

$$\Delta \ln s_{gct} = \varphi_{gt} - (\sigma_g - 1)\Delta \ln p_{gct} + \varepsilon_{gct}$$

where φ_{gt} is a function of the same variables that enter the price index for good g and does not differ across the country of origin (it is a random effect in the demand equation). $\varepsilon_{gct} = \Delta \ln d_{gct}$ is the random term in the regression and can be interpreted as taste or quality shocks across import destination.

The export supply equation is specified exogenously, but allowed to vary with the amount of exports:

$$\Delta \ln p_{gct} = \psi_{gt} + \frac{\omega_g}{1 + \omega_g} \Delta \ln s_{gct} + \eta_{gct}$$

where $\varphi_{gt} = -\omega_g \Delta \ln E_{gt} / (1 + \omega_g)$; $\omega_g \geq 0$ is the inverse supply elasticity, assumed to be constant across countries and $\eta_{gct} = \Delta \ln v_{gct} / (1 + \omega_g)$ captures any random change in a technology factor v_{gct}. A special case of this supply equation would be a horizontal, perfectly elastic supply, in which case $\omega_g = 0$.

Obviously, the system of demand and supply is not identified without instruments or identifying restrictions. Lacking instruments for the entire range of industries we would like to estimate this system for, we instead assume that $E(\varepsilon_{gct} \eta_{gct}) = 0$, i.e. once good-time specific effects are controlled for, demand and supply errors at the variety (destination country) level are assumed to be uncorrelated.

Details of the estimation procedure follow Leamer (1981)[51]. Both equations are normalized by a reference variety (k) and parameters are estimated from the second moments in the data. The estimation equation becomes:

$$(\Delta^k \ln p_{gct})^2 = \theta_1 (\Delta^k \ln s_{gct})^2 + \theta_2 (\Delta^k \ln p_{gct} \Delta^k \ln s_{gct}) + u_{gct}$$

where

$$\theta_1 = \frac{\omega_g}{(1 + \omega_g)(\sigma_g - 1)}, \quad \theta_2 = \frac{(1 - \omega_g)(\sigma_g - 2)}{(1 + \omega_g)(\sigma_g - 1)}, \quad \text{and} \quad u_{gct} = \varepsilon_{gct}^k \eta_{gct}^k$$

All the k superscripts indicate differencing by the corresponding variable for the kth reference country. In order to recover the structural parameters of interest σ_g and ω_g, we have to solve the nonlinear system of equations. For some parameter values, there will only be imaginary solutions, which obviously do not make economic sense. The endogeneity of the price on the right-hand side is solved by instrumenting with country dummies, see Feenstra (1994) for details on the practical implementation of the weighted IV estimator.

---

[51] Leamer, E. (1981), "Is it a Demand Curve, or Is It a Supply Curve? Partial Identification through Inequality Constraints," *Review of Economics and Statistics*, 63(3), pp. 319-327.

To assess the reasonableness of these parameter estimations we present our estimates for the σ_g parameters in 4.10, side by side with the estimates obtained by Broda and Weinstein (2006). We follow these authors by including different goods in the same SITC 3-digit sector imported from one country as different varieties. We present estimates for all automotive sectors (broadly defined). The first two columns replicate the demand elasticities that Broda and Weinstein (2006) obtained using a GMM estimator that avoids the risk of finding imaginary values for any of the structural parameters by doing a grid search restricted to the allowable interval. In the next column we present our results using the same data, made available by Feenstra[52]. In the fourth column, we replicate the analysis on a more recent data set we purchased from Global Trade and Information Services, Inc. The latter spans the 1995-2005 period.

Table 4.10. Demand elasticity estimates at the SITC 3-digit level

SITC (Revision 3)		Broda & Weinstein		Replication	GTIS data:'95-'05	
		('72-'88)	('90-'01)	('90-'01)	(1)	(2)
621	materials of rubber	3.52	2.67	1.77	2.31	2.29
625	rubber tires	2.18	1.98	3.70	4.96	3.10
713	**internal combustion engine**	**25.03**	**2.69**	**3.46**	**3.21**	**2.42**
781	motor vehicles for passengers	1.63	3.02	15.55	2.21	2.48
782	motor vehicles for goods	103.03	6.70	5.00	9.46	N/A
783	road motor vehicles, n.e.c.	10.59	3.82	7.71	367.44	N/A
784	**parts and accessories for MV**	**7.76**	**2.79**	**2.04**	**1.48**	**1.68**
785	motorcycles and bicycles	1.30	1.15	2.52	1.83	2.08
786	trailers	3.73	1.89	2.00	1.77	1.87
Nobs		246	256	9	9	7
Mean		6.78	4.01	4.86	43.85	2.28
St.Dev (Mean)		1.19	0.50	1.48	40.46	0.18
Median		2.54	2.24	3.46	2.31	2.29
Max		1.08	1.05	1.77	1.48	1.68
Max		228.75	108.19	15.55	367.44	3.10

Notes: The first three column use data available online through the NBER website. The last two columns replicate the same analysis using more recent data from the Global Trade Atlas, compiled by Global Trade Information Services, Inc. (1): normalization by largest Canadian hs10 product, in terms of trade volume; (2): normalization by smallest Canadian hs10 product.

In each estimation, market shares are calculated using reported import values for the U.S., omitting re-exported goods. Prices are constructed as unit values, dividing the value of imports by quantity. Ideally, we would like to carry out the estimation for Canada as well, but for the automotive parts, all goods starting with 8708 in the Harmonized System of trade classification had missing quantities.

Simply comparing the results in the first two columns, the Broda and Weinstein (2006) results already indicate that esti-mates are not always robust over time. For example, motor vehicles for goods transport have an estimated elasticity of 103 in the early period 1972 to 1988, basically indicating these are homogenous goods. In the later period, 1990 to 2001, the elas-ticity declines to 6.70 indicating some differentiation—optimal monopoly mark-ups would increase from less than 1% to 17.5%. While it is not entirely impossible that these goods have changed this much over time, it is doubtful. For some other goods, the parameters are much more stable over time. For ex-ample, the demand for rubber tires and motorcycles & bicycles is estimated to be highly inelastic in both periods.

For some goods our results are relatively close to the ones obtained by Broda and Weinstein (2006), but this is not always the case. Most pertinent for this section are the estimates for the internal combustion engine & parts sector (SITC 713) and mo-tor vehicle parts & accessories (SITC 784), which are both indicated in bold. Our results are relatively close. For a lot of goods, Broda and Weinstein (2006) find a declining trend in the elasticities, indicating increased product differentiation over time. We find similar results and extending the data set to 2005, results in the fourth column, tends to lower the elasticities further.

Because the estimation procedure is nonlinear, the results might be sensitive to the normalization chosen. We have always normalized by a Canadian good, because the U.S. has positive imports from Canada in the largest set of products throughout the entire sample period and also in terms of market share Canada is important and stable over time. However, at this aggregate level we lump a lot of products (at the most detailed 10-digit HS classification) together when we carry out the esti-

mation for each SITC 3-digit industry separately. In the fourth column, we normalize by the Canadian product with the largest market share, while in the fifth column the estimation procedure is identical, only now we normalize by the Canadian product with lowest market share. Results are somewhat sensitive to this normalization, but there is no consistent direction for the bias. Also note that with this alternative normalization we cannot obtain the demand elasticities for industries 782 and 783.

To gauge whether the results are still reasonable when we carry out the estimates at finer levels of aggregation, Table 4.11 contains the demand elasticities for the detailed sub-sectors in engines & engine parts (SITC 713) and automotive parts and accessories (SITC 784). In the first two rows we repeat our estimates from the 3rd and 4th column of Table 4.10, for the aggregate sectors, and we add the estimates of the supply elasticities as well. One pattern that seems to come out of this is that goods with a high demand elasticity tend to have a lower supply elasticity. Combinations of high demand and supply elasticities, which would give rise to large quantity volatility over time, seem to be rare. Similarly, we only find a single good with a demand elasticity below the median that also has a supply elasticity below the median (motor vehicle bodies, SITC 78421). A situation like this is likely to lead to high price volatility over time when either of the curves shifts.

Overall, the estimates seem reasonable. The median demand elasticity across all parts sectors is 2.53 in the 1990-2001 period and 2.70 in the 1995-2005 period. The averages are larger because some sectors are estimated to have much lower product differentiation, while the estimate for the demand elasticity can never fall below 1 (to be consistent with the model). Median supply elasticity is 0.78 in the 1990-2001 period, notably below infinity (the perfectly competitive benchmark). This makes sense: as we estimate that products are differentiated, it makes sense to find that the supply curve is not entirely horizontal. We also find that the supply elasticity declined in the later period. The median declined to 0.36 and, omitting the outlier motor vehicle engines with spark-plugs larger than 1000cc (SITC 71322), the average supply elasticity also declined from

0.82 to 0.65. Increased product differentiation seems to make it harder for firms to quickly scale up production. Increased foreign competition would tend to increase the supply elasticity.

Table 4.11. Demand and supply elasticity estimates at the SITC 5-digit level

	Feenstra data(1990-2001)		GTIS data (1995-2005)		
	demand	supply	demand	supply	description
713	3.46	0.33	3.21	0.38	
784	2.04	0.97	1.48	1.61	
71321	(b)	(b)	(b)	(b)	MV engine – spark < 1L
71322	2.59	2.48	2.31	19.8	MV engine – spark > 1L
71323	(a)	(a)	8.83	-0.3	MV engine – diesel
71381	8.71	0.02	6.97	0.75	engine other, spark
71382	7.36	0.14	(a)	(a)	engine other, diesel
71391	2.52	0.78	2.03	0.97	parts – spark
71392	2.27	0.41	2.51	0.35	parts – diesel
78410	(a)	(a)	36.92	0.35	chassis with engine
78421	2.06	1.44	(b)	(b)	MV bodies
78425	(a)	(a)	9.57	0.21	other bodies
78431	2.38	0.81	2.9	-0.15	bumpers
78432	1.6	0.86	1.64	0.95	other parts of bodies
78433	4.31	-0.17	2.3	0.29	brakes
78434	1.94	3.32	1.95	3.38	gearboxes
78435	3.14	2.27	6.07	0.01	drive-axles
78436	2.86	-0.23	5.23	0.38	non-driving axles
78439	2.53	0.21	1.63	1.22	other MV parts
Nobs	**13**	**13**	**14**	**14**	
Mean	**3.41**	**0.95**	**6.49**	**2.01**	
St.Dev (Mean)	**0.6**	**0.31**	**2.45**	**1.39**	
Median	**2.53**	**0.78**	**2.7**	**0.36**	
Min	**1.6**	**-0.23**	**1.63**	**-0.3**	
Max	**8.71**	**3.32**	**36.92**	**19.8**	

Notes: (a) imaginary number; (b) no data

Finally, in Table 4.12, we also show the demand and supply elasticities estimated at the 6-digit level of aggregation for the Harmonized System of trade classification. This is the level of detail that we will use to simulate the impact of the changed trade policy. The same products as in Table 4.11 are included.

Table 4.12. Demand and supply elasticities at the 6-digit HS classification

Harmonized System	Feenstra data (1990-2001)		GTIS data (1995-2005)	
	Demand	Supply	Demand	supply
8407-8409	3.53	1.09	3.16	0.81
8708	2.03	0.59	1.47	1.30
840730	2.07	-2.69	2.56	303.74
840731	3.91	-0.47	2.00	2.52
840732	6.11	0.75	10.08	1.36
840733	4.48	0.31	26.49	1.11
840734	2.59	2.48	2.31	19.80
840790	8.71	0.02	6.97	0.75
840820	(a)	(a)	8.53	-0.30
840890	7.36	0.14	(a)	(a)
840991	2.88	1.28	1.89	4.91
840992	2.81	0.77	2.31	0.64
840999	2.25	1.29	2.51	0.35
870810	2.38	0.81	2.90	-0.15
870820	2.96	1.75	(b)	(b)
870821	(b)	(b)	4.80	0.02
870829	1.55	0.44	1.64	0.95
870831	4.60	2.05	1.85	3.44
870839	3.22	-0.10	3.22	-0.52
870840	1.94	3.32	1.95	3.38
870850	3.14	2.27	6.07	0.01
870860	2.86	-0.23	5.23	0.38
870870	3.38	-0.20	1.85	1.43
870880	1.77	3.30	2.82	0.39
870891	1.84	-7.11	1.80	5.17
870892	2.37	1.05	1.53	2.03
870893	2.14	0.17	1.43	4.18
870894	2.04	0.16	1.89	-0.24
870899	2.29	0.02	1.87	0.84
HS 6-digit				
Nobs	27	27	27	27
Estimated	25	25	25	25
# of varieties (median)	40	40	39	39
Mean	3.27	0.46	4.26	14.25
St.Dev (Mean)	0.36	0.41	1.04	12.09
Median	2.81	0.44	2.31	0.95
Min	1.55	-7.11	1.43	-0.52
Max	8.71	3.32	26.49	303.74

Notes: (a) imaginary number; (b) no data

With the exception of a few outliers, the estimates are now very similar for the two time period. Engines & engine parts, those components starting with 8407, 8408, or 8409 seem less differentiated than other automotive parts. They have larger demand elasticities, although the supply elasticity is only estimated to be larger in the earlier period.

4.5.2 Export potential

Using the U.S. estimates in the previous section as indicative of the demand and supply estimates for all exporters in all markets, we can simulate what the impact of a trade agreement would be on Canadian exports. While it might seem like a strong assumption to use the U.S. estimates for other countries, it is not that farfetched in this case. The production technology used in the automotive industry is the same the world over. The same firms are also operating assembly plants in all of the regions we will investigate and in the U.S. This should make the demand elasticity estimates—which are input factor demands—comparable. Moreover, the Canadian industry is the most important trading partner of the U.S. and supply elasticities identified from U.S. imports should be highly representative of the Canadian industry. As such, the supply elasticities should be equally valid. Note that for the estimates in the previous section to be valid, we had to assume anyway that supply and demand elasticities were identical across countries. If countries differ substantially in technology, this will show up in the relative importance of different goods in their trade flows[53].

To calculate the impact of an elimination of tariff rates under FTAs between Canada and different countries, we exploit the properties of the CES demand system that underlies our

---

[53] As mentioned earlier, we were unable to carry out a similar estimation for Canada, as quantity data was not available. For the U.S. trade flows, the Feenstra data has missing observations for physical quantity (import values are always available if trade flows are positive) in less than 10% of the observations. In the Trade Analyzer Database of Statistics Canada, the automotive parts information had more than 95% of the quantities missing (finished vehicles were reported in physical units).

elasticity estimations. A good reference for the crucial equations is Melitz (2003)[54]. In the CES model, mark-ups will be constant for all producers as each firm/country faces a residual demand curve with the same elasticity of substitution (which differs by product). Price will be set by the following mark-up pricing rule:

$$p_{ct} = c_{ct}(1+t)\frac{\sigma}{\sigma-1},$$

where c_{ct} is the marginal cost of production (including transportation) for country c at time t, existing import tariffs increase the marginal cost and the mark-up is only a function of the product's elasticity of demand in the importer country. For the moment, we assume marginal costs are constant, i.e., the supply curve is perfectly elastic (this will overestimate the effect of the trade policy change). We relax this assumption later.

Total imports for each country are given by

$$m_{ct} = R_t\left(\frac{p_{ct}}{P_t}\right)^{1-\sigma},$$

where the relative price is what matters, i.e., country c's price relative to the aggregate price index for the importing country. Given that Canada is a relatively small trading partner for the countries we consider, we will throughout take total import spending R and the aggregate price index P to be exogenous to Canada's price and quantity choices. In this case it is straightforward to derive that the impact of a trade policy on Canadian exports will be given by

$$\frac{\partial \ln m_{ct}}{\partial t} = (1-\sigma)\frac{\partial \ln p_{ct}}{\partial t}.$$

[54] Melitz, M. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity, *Econometrica*, 71(6), November, pp. 1695-1725.

Given the constant mark-up in the pricing rule, the derivative on the right-hand side is equal to $\Delta t/(1+t)$. The model predicts that the mark-up will not change and that the elimination of the tariff will be passed on to consumers proportionally to the current final good price. At the same time, the quantity sold will increase by $(-\sigma)$ times the price decline. Note that the CES model of monopolistic competition assumes that Canadian firms do not take the effect of their pricing decisions on their competitors' behaviour into account. Given the low market share of Canadian firms (between 1% and 3% in the different markets), this seems like a plausible assumption.

In our calculations we take into account that different products face very different demand elasticities, as calculated in the previous section. Hence, the distribution of Canadian exports over the different sub-sectors will be very important. We calculate the effect of three possible free trade agreements (FTAs)—with China, with South Korea, and with the enlarged E.U. (with 25 members)—under which Canadian parts exporters would see their import tariffs eliminated. We do not include an FTA with Japan, as it currently does not impose any import tariffs on automotive parts.

Current Chinese tariff rates are 15%, but they are scheduled to decline to 10.3% under its WTO agreement. South Korean tariffs are, as far as we have been able to determine, currently 8% uniformly across all automotive parts. European tariffs are lower and vary somewhat by category. On small engines and most engine parts, the E.U. levies 2.7%, on larger engines the rate is 4.2%. On most other components it levies 3% duties, but on a whole range of "not elsewhere classified" products, the duty is 4.5%. For China and the E.U. we present two alternative sets of results, using the lower or the upper range for the tariff rates.

For all three countries, we also present alternative calculations using the highest demand elasticity estimate obtained for each good, see Table 4.12, or using the low demand elasticity. We calculate the absolute value of the expected import increase (Canadian exports) by summing over all the 6-digit HS categories. We also express the total amounts as a fraction of the current Canadian export levels. The top panel of Table 4.13

contains the four sets of results if we use the 2005 trade flows as benchmark, the bottom panel uses the average 2004-2005 trade flows as a robustness check to guard against annual fluctuations.

Table 4.13. Counterfactual simulations of Canadian import changes following separate FTAs with three trading partners

demand elasticity	initial tariff	China		South Korea		E.U. 25	
2005 Canadian import levels							
min	low	$8.61	10.4%	$1.43	8.4%	$5.90	3.4%
min	high	$12.03	14.6%			$8.83	5.4%
max	low	$13.13	15.9%	$1.98	11.6%	$9.08	5.2%
max	high	$18.34	22.2%			$13.58	7.9%
Average 2004-2005 Canadian import levels							
min	low	$17.89	10.2%	$0.96	8.7%	$5.33	3.7%
min	high	$25.00	14.3%			$8.00	5.5%
max	low	$26.88	15.4%	$1.34	12.2%	$8.05	5.6%
max	high	$37.54	21.5%			$12.07	8.3%

Note: Effects are expressed in million CDN$ or as a percentage of initial parts exports

Before we discuss the results, it is important to stress that this is a counterfactual analysis, not a prediction. We tried to assess what Canadian exports would have been had exports not been subject to import tariffs in 2005. Given the rapid change in the industry, this is quite different from the expected export change if tariffs will be eliminated in the future. Even without any trade policy change the 2006 statistics are likely to look very different from 2005[55].

The predicted changes are rather large. For 2005, they range from 10.4% to 22.2% for a FTA with China, from 8.4% to 11.6% for a FTA with South Korea, and from 3.4% to 7.9% for a FTA with Europe. Given that the average demand elastic-ity is around two (using the "min" elasticity), we find that quantity exported increases approximately twice as much as

[55] As an example, Canadian parts exports to China nearly halved from 2003 to 2004 as shipments of body panel for the Buick model assembled in Shanghai stopped.

price decreases and the net gain in import revenue is roughly equal to the initial tariff rate. When we use higher demand elasticities, the price response does not change, but quantity changes more, leading to larger effects.

The difference in import response across the different trading partners is predominantly the result of differences in the current tariff rates. However, the composition matters as well. For example, exports to the E.U. are largest in the HS 870829 category, which has one of the lowest demand elasticities. As a result, total exports to the E.U. increase slightly less than the average tariff decline, while we find the reverse for the FTAs with China or Korea (a slightly more than proportional export increase). Similarly, for Europe the predicted import increase of Canadian products is almost twice the average level of tariff reduction in the most optimistic case, reflecting that for some important goods (in particular HS 870870 and 840891) the high demand elasticity is almost double the low elasticity. High elasticities mean low mark-ups and the same price reduction has more impact and, moreover, consumers are more responsive.

Another point worth noting is that the absolute values are sizeable. Even though China's automotive industry is not very large yet, it does import a lot of components. This tends to be typical of low-wage assembly centres. Imports of components outstrip exports and the reverse holds for finished goods. In the automotive industry, China is a large importer of both components and finished vehicles. The large import response we estimate would translate into extra sales for the Canadian industry in China with estimates ranging from $8.79 to $18.34 million CDN, depending on the choice of demand elasticity and initial tariff level. Consistent with the previous discussion, we find smaller effects in the most recent year (using the 2005 import data) than using average imports for the 2004-2005 period. It indicates that the Chinese industry is rapidly increasing its level of self-sufficiency.

Finally, we cannot distinguish between OE parts and aftermarket parts in this part of the analysis. The trade statistics are not broken down along that dimension. We do not know the share of aftermarket parts in total parts exports, but it is likely to

exceed the share of aftermarket parts in domestic production. While Canadian firms are very tightly integrated in the North American industry, that is much less the case overseas. Canadian exporters often claim they have a very hard time penetrating overseas OEM markets. Given that Japanese and Korean firms take a long time to increase the domestic content in the vehicles they assembly in North America, this seems plausible. The current analysis assumes all Canadian parts exports are aftermarket, which is obviously an overestimate of the likely effect of the FTAs, but not necessarily a very large overestimate. For the actual effect, we should simply pro-rate the absolute import increases (by the share of aftermarket parts in exports), while the percentage changes are valid.

Another factor that would lower the expected impact of tariff reductions is an upward sloping supply curve. The results in Table 4.14 incorporate the estimated slope of the supply curve in the simulations. The estimated supply elasticities correspond to the inverse of the effect of increased output on marginal costs. Incorporating this effect in the pricing rules gives

$$\frac{\partial \ln p_{ct}}{\partial t} = \underbrace{\frac{\partial \ln c_{ct}}{\partial t}}_{=\frac{\partial \ln c_{ct}}{\partial \ln q_{ct}}\frac{\partial \ln q_{ct}}{\partial t}} + \underbrace{\frac{\partial \ln(1+t)}{\partial t}}_{=1/(1+t)} + \underbrace{\frac{\partial \ln(\sigma/(1-\sigma))}{\partial t}}_{=0} .$$

Table 4.14. Counterfactual FTA simulations with an upward sloping supply curve

demand elasticity	supply elasticity	initial tariff	China		South Korea		E.U. 25	
2005 Canadian import levels								
mean	mean	low	$2.68	3.25%	$0.31	1.83%	$1.84	1.07%
mean	mean	high	$3.75	4.54%	$0.31	1.83%	$2.76	1.61%
max	max	high	$5.20	6.30%	$0.54	3.14%	$4.36	2.54%

Note: Effects are expressed in million CDN$ or as a percentage of initial parts exports

The first term under the first bracket is $(1/\omega)$, the inverse of the supply elasticity, the second term can be expressed as

a function of the price change by taking the derivative of the relative import demand, which gives $-\sigma \dfrac{\partial \ln p_{ct}}{\partial t}$. Solving for the optimal price response, we find that the earlier price change $\Delta t/(1+t)$ is adjusted by a factor $(1 + \sigma/\omega)^{-1}$. A higher supply elasticity (ω) will lead to a larger price response as marginal costs are almost constant. A higher demand elasticity (σ) will lead to a smaller price decrease for any positive supply elasticity, because it leads to larger cost increases as quantity is very responsive.

When we perform the counterfactual simulations taking supply effects into account, the import effects are much smaller, not surprisingly given the low supply elasticities estimated in the previous section. In the first line of Table 4.14 we use the average demand and supply estimates and the low initial tariff rates. In the next line, we use the high tariff rates instead. In the third line, we calculate the most optimistic FTA effect if supply effects are taken into account. We take the maximum estimate for the demand elasticity (which makes quantity very responsive to price declines), the maximum estimate for the supply elasticity (to minimize the increase in marginal cost), and we take the high initial tariffs. Even under this scenario, the estimated impact is reduced by a factor of more than 3 in each country, relative to the upper range of the predictions in Table 4.13. As before compositional effects are important. Even though the tariff reduction is almost twice as high in Korea compared to Europe, the composition of Canadian exports leads to a comparable percentage effect in both countries/regions.

A final caveat is that these calculations assume that the impact of a free trade agreement is proportional to the decline in tariff rates. There are some indications that non-tariff barriers are important. While this is particularly true for final vehicles, it probably affects parts and accessories as well[56]. Unfortunately, there is no clear way to quantify this.

---

[56] Even though the Japanese and Korean market shares in the U.S. and Canada vehicle markets are sizeable by now, they have not been able to

4.5.3 Import competition

Even though the calculations would be straightforward, we refrain from carrying out the same type of analysis as in the previous section to study increased import competition in Canada. We expect it would result in very implausible estimates because the large domestic automotive assembly sector imports a lot of OE parts that already enter duty free. For example, while Canada exported $29.2 million worth of parts to Japan in 2004, imports stood at $1743.2 for a deficit of $1713.9 millions. Obviously the majority of these parts enters the assembly process and do not incur any duties. Given the small size of the aftermarket, discussed earlier, especially relative to the large volume of imports, any estimate of the fraction of imports going to the aftermarket would be subject to a very large margin of error.

With South Korea, the E.U., and China, the Canadian deficit in parts is also very large, running to $266.0, $636.3, and $664.7 million respectively, in 2004. Including trade with the US and Mexico, total parts imports in Canada were $42,859 million in 2004 which was approximately 10 times larger than the entire aftermarket parts sales.

Given that in Table 4.10 the rubber tires and parts & accessories industries—the two sectors containing the bulk of aftermarket parts—were estimated to have the lowest demand elasticities of all automotive industries, import responses are likely to be moderate. The effects on import values abroad we estimated in the previous section are a combination of price declines and quantity increases. In sectors with a lot of product differentiation (low demand elasticity) as the aftermarket parts sector, quantity increases will be low– less harm to the Canadian industry, but these will be accompanied by more moderate price declines—smaller benefits for the Canadian consumers.

---

penetrate each other's market. In 2003, only 3774 Japanese passenger cars were sold in Korea and 2573 Korean cars in Japan, which represent a market share of 0.28% and 0.04% respectively.

4.6   Pricing-to-market

In section 4.5, the simulations are carried out assuming all firms behave competitively, i.e. the CES demand system leads to a monopolistically competitive industry, where the slope of the residual supply function is not affected by a firm's actions or any action of its competitors. In particular, we have assumed that Canadian exporters do not take the response of foreign competitors into account when they decide how much to lower their price following a tariff cut. While this is probably a good assumption given the small Canadian market share overseas, we verify here how sensible this assumption is.

Using the methodology developed by Goldberg and Knetter (1999)[57] we estimate the slope of the residual demand Canadian firms face in the markets abroad considered in the previous Section. The idea is to identify the slope of the residual demand exploiting exchange rate variability as an indicator for cost changes. Note that these will differ from the demand elasticities estimated earlier. Residual demand elasticities include supply responses by competitors, which will depend on the type of market equilibrium the industry is in—which is not specified explicitly. The estimating equation is as follows:

$$\ln p_{mt}^{ex} = \lambda_m + \eta_m \ln Q_{mt}^{ex} + \alpha' \ln Z_{mt} + \beta' \ln W_{mt} + \varepsilon_{mt}$$

The residual demand for exporter ex, to destination market m, at time t, expresses the price the exporter charges (in the importing market's currency) as a function of its own quantity, demand shifters for the overseas market (Z) and cost shifters for its competitors (W).

Given that we only have 11 years of data to estimate the equation, we have to be extremely parsimonious in the specification. The only demand shifter we include is a time trend and as cost shifters we use the exchange rate of the two largest importers, apart from the country under investigation. As

---

[57] Goldberg, P. K and M. M. Knetter (1999), "Measuring the intensity of competition in export markets", *Journal of International Economics*, 47, pp. 27-60.

instrument for the endogenous quantity level, we use the exporter's exchange rate—a valid cost shifter[58]. The results in Table 4.15 are obtained from separate regressions for each of the three regions and for each of the five exporters. Even though the countries sell in the same import market, they will face a different residual demand, because they face different competitors.

Table 4.15. Elasticities of the residual demand curve for total automotive parts

	South Korea		China		E.U. (25)	
	market share	residual demand elasticity	market share	residual demand elasticity	market share	residual demand elasticity
Canada	**0.6%**	**0.184** (0.63)	**1.1%**	**-0.354** (2.12)	**1.8%**	**-0.710** (2.88)
Japan	39.3%	-1.052	42.5%	-11.990	53.8%	-1.396
United States	15.9%	-1.464	8.1%	0.078	32.5%	-0.758
Euro-area	34.3%	-2.667	22.7%	-1.354		
China	5.5%	0.188			6.1%	-0.134
South Korea			22.0%	-0.743	3.1%	0.632

Note: Sample includes all imports of engines & engine parts and automotive parts and accessories over the 1995-2005 period. t-statistics in parenthesis

Even though Canada has a very small market share in each of the three markets, we find that it has a surprising amount of market power. The statistics are the inverse of the elasticities, as is customary in this literature. An estimate of -0.71 for Canadian exports to the E.U. corresponds to a residual demand elasticity of only 1.41. In the Chinese market, Canada is still estimated to have a decent amount of market power, with an implied elasticity of 2.83, significantly different from a perfectly elastic residual demand. Only in Korea, we find no

market power for Canadian firms. The results are relatively similar once we limit the estimation to only a single sub-sector of automotive parts, but many parameters become unstable.

While Canada has some market power, the implied residual demand elasticities we find for Japan, the E.U., or the U.S. are notably lower. Only in the Chinese market do Canadian firms have more pricing power than the U.S. In contrast, Chinese firms are never estimated to have any significant market power and the only negative coefficient for South Korea, for its exports to China, is estimated very imprecisely (t-statistic is 1.21)—even though its market share is quite large. These results are intuitive as we would expect the countries with the most developed automotive industries to have the most sophisticated and differentiated goods, and hence the largest market power—which lines up well with the estimates in Table 4.15.

In light of these results and the very low market share for Canada in automotive parts exports to South Korea, China, and the E.U., we think that it is plausible to assume Canadian firms will not act strategically in their response to tariff policy changes, as was assumed in Section 4.4.

5.   Future direction of the industry

This section analyzes the future direction of the auto industry, particularly in the North American market, in the next 5-20 years, Canada's potential to move towards high-value production of auto products, and the potential to attract future assembly and production of auto production. What would be needed, including vis-à-vis trade policy, to promote high-value production?

5.1   Future direction

Most of the relevant issues have already been addressed in the preceding Sections. Here I just summarize the most important trends—most important in terms of likely future impact.

5.1.1   Fuel

The great unknown for the industry is what type of fuel cars will drive in the future. Currently, the vast majority of vehicles today use a gasoline internal combustion engine, but that is likely to change in the not so distant future[59]. The corporate average fuel efficiency norms (CAFE) have been tightened repeatedly for cars, and the Bush Administration finally raised the standards for light trucks as well, which currently account for more than 50% of new vehicle sales in the U.S.[60]. More efficient, direct injection gasoline engines have started to appear, but more radical alternatives are also on the horizon.

It is not impossible that diesel engines will become much more popular in North America. In Europe they already account for more than 50% of new vehicle sales. A number of big manufacturers, especially Volkswagen and DaimlerChrysler, are committed to offer a greater selection of diesel engines in their passenger vehicle lineup. The advent of clean (low-sulfur) diesel

---

[59] The most popular current alternatives are diesel (especially for pickup trucks), LPG (especially for taxis and limos), and hybrids.

[60] Automotive News, March 29, 2006, "Fuel economy is toughened for 2008-2011 trucks".

in the U.S. in 2006-2007 is a pre-requisite for modern diesel engines. The future of diesel is by no means secure. The cost penalty to lower diesel emissions to the same level of gasoline emissions is proving to be much more costly than anticipated. At the same time, many of the fuel-saving technological advances common in today's diesels are often not yet introduced in gasoline engines (variable valve timing, turbo charging, and direct injection). Fiat estimates that a diesel engine would cost 1,000 euro more in production than a comparable gasoline engine, have the same CO_2 emissions, and only a 5-10% fuel efficiency advantage. A U.S. government plan to provide tax incentives for diesel could prove decisive, but the details remain to be determined. Currently, no diesel engines are manufactured in Canada.

Furthermore, another advantage of the gasoline engine is that it can burn alternative fuels, such as natural gas or hydrogen, with few modifications. This multi-fuel use has proved to be a boon for the electricity generation industry, where modern generating stations switch between natural gas or oil depending on the price of the month. A diesel engine cannot achieve the same feat, although several European countries, especially Sweden and Germany, are making bio-diesel, made from organic material widely available.

While the engine is the most expensive single part of a vehicle, the alternative powertrains considered would have a much greater impact on the structure of the industry. Foremost, hybrids are likely to keep increasing in popularity. Total hybrid sales in 2005 exceeded expectation at 205,749 in the U.S. alone—52% of this by the Toyota Prius, representing 1.2% of all new vehicle sales. This ranges from a share of hybrids in total sales of approximately 8% for Toyota, 4% for Honda, less than 1% for Ford, and negligible for all other manufacturers[61]. A survey of North American automotive executives by KPMG yielded a unanimous prediction that hybrids will increase market share in the coming years[62]. Currently, no hybrids are

---

[61] GM has delivered 430 diesel-hybrid busses in the U.S. and Canada and expects to add 237 to that total in the remainder of 2006.

[62] On the sample of worldwide automotive executives, 88% anticipated a rising market share for hybrids.

manufactured in Canada. This will change once Ford introduces the hybrid versions of the Ford Edge and Lincoln MKX in its Oakville assembly plant. The gasoline versions will start production in 2006 as a 2007 model, but production of hybrids is expected to start only in 2010.

Plug-in hybrids could also become more popular in the future. These are hybrid cars with an enlarged battery pack that can be recharged from the electricity grid, not only by the onboard gasoline engine. For the vast majority of trips, only the electric engine would be used and the battery pack recharged overnight or at the office. Only on longer trips would the combustion engine be used. This setup does away with a major disadvantage of the previous generation of electric cars: the risk of getting stranded if the battery runs out.

Much further down the line is the changeover to the hydrogen economy and vehicles driven by fuel cells. Current expectations of most automakers are that by 2010 most of the technical aspects will be solved on the experimental models that are now touring the globe. It is also expected to take until 2020 or so before mass manufacturing would make affordable cars possible. An average sized car currently can store about 3 kilograms of pressurized hydrogen gas which can go about 200 to 280 kilometres under normal conditions before refuelling. Developing reliable storage for hydrogen and rolling out a distribution system are considered the biggest challenges for this new technology. As discussed in Section 1.5, Canada is very active in the development of fuel cells.

5.1.2   Assembly location

The second great unknown for the industry is whether final vehicle assembly will stay as close to customers as it has thus far. In the first decades after World War II, the industry produced very large production runs of a small number of vehicles in branch assembly plants close to population centres. For example, U.S. sales of the different guises of the main Chevrolet model totalled almost 1,500,000 units in 1966 and these were assembled in six different assembly plants across the country.

Declining model runs have resulted in very few cars or light trucks being produced in more than one U.S. plant. Except for a few instances where models are moved between assembly plants, not a single car was produced in more than one U.S. assembly plant in 2004[63]. The larger average production run for light trucks makes it more common for them to be assembled in more than one location, but the recent proliferation of crossover vehicles is lowering production volumes of light trucks as well.

Firms could have decided to develop vehicles for the global market, produce them in a single country, and ship them around the world, as is the current practice in the consumer electronics industry. With a few exceptions, this has not happened. Instead firms are investing in flexible manufacturing systems in order to build multiple vehicles on each assembly line[64]. This allows firms to produce a wide range of vehicles on each continent. It is unlikely that this decision is to a large extent driven by trade policy. Most developed countries charge only modest import duties on vehicles. Shipping costs for a bulky and easily damaged (scratched) product like an automobile are likely to be non-negligible and not decreasing over time. Proximity to consumers in a mature industry, where responding quickly to changing tastes is important, is probably another important factor.

The industry has also repeatedly flirted with made-to-order systems. The current industry benchmark for new vehicle inventories in the U.S. is 60 days, which is worth at least $60 billion (US). That is a lot of working capital sitting idle. In practice, inventory is larger for most vehicles; surprisingly, inventories tend to be higher for domestically produced vehicles. In a market where the number of available varieties totals almost 300, the risk of mismatching production and sales is enormous. The potential cost, in terms of a forced discount, to be able to sell undesirable vehicles, is corre-

---

[63] A number of the largest volume vehicles, such as the Toyota Corolla, Chevrolet Malibu, or Ford Focus, are still produced in more than one North American plant.

[64] See Van Biesebroeck (2006), "Complementarities in Automobile Production", NBER Working Paper for a discussion.

spondingly enormous. Made-to-order, which is popular in Europe, would tie the assembly location to North America.

Even though a lot of components are currently outsourced to low cost countries, this has not happened with finished vehicles. Two Chinese firms are planning to start exporting finished vehicles from China to North America (Gheely in 2007 and Chery in 2008) and Honda is exporting its Fit subcompact car from Guangzhou to Europe. While the labour cost in assembly is too small a fraction of the total cost of a cart to justify producing it in a low-wage country like China, the trade-off changes once firms would be able to leverage the lower wages over the entire supply chain and produce most components at the low Chinese wage as well. The rapid development of the Chinese domestic industry is quickly making this a possibility. For example, the engines installed in the Chevrolet Equinox in the CAMI plant are shipped from China.

It took Japanese producers only 10 years after their first sales success in North America to establish local assembly capacity, although this choice was accelerated by the voluntary export restraints. Hyundai did not even wait this long to (entirely voluntarily) open up its first assembly plant in Alabama; its second plant, for its Kia subsidiary, has been announced at a time when it is not even certain it will be able to operate the plant at full capacity[65]. An important distinction is that the much larger labour pool in China is likely to keep wages depressed for a longer time than in Japan or Korea. It might make China an attractive assembly location for exports to North America in the future. The major Chinese car and component producers are currently benefiting a lot from their collaboration with leading western automotive companies. They are unlikely to anger their joint venture partners by challenging them in their home market. However, once they feel they have learned what they wanted to know, their incentives will change.

A final issue that has come up repeatedly in this report is the future division of labour between OEMs and suppliers. Many tier 1 suppliers are playing an increasing role in R&D and

---

[65] An earlier venture in Quebec in 1989 was of a much lower scale.

329

design. Several firms, e.g., Magna International, are also taking on assembly tasks. Whether this makes it less or more likely that assembly will move overseas remains to be seen.

5.1.3   Volume

As discussed at length in section 1.4, we anticipate a reduction in future North American sales (and production), especially in the U.S. Registrations are at an all-time high and especially very new vehicles abound after a sustained string of bumper sales years. Many industry observers rationalized the large Big 3 sales decline in the fall of 2005 as a mismatch between the gas-guzzling vehicles they produce and consumers' newly acquired taste for fuel-efficiency in the post-hurricane Katrina spike in fuel prices. It seems much more likely that the spectacular summer sales, fuelled by employee discount programs, were responsible. Over the summer, large SUVs had been the most successful market segment and the market is probably saturated[66].

The increased durability of modern vehicles will make sure the current stock of vehicles will be around for quite some time. The large number of fuel-inefficient SUVs and other types of trucks sold in recent years make the fleet of second-hand vehicles less suitable for exporting to less developed economies, especially with the current high fuel price. Exporting new vehicles to keep assembly plants operating at full capacity seems also a very unlikely proposition, as discussed in section 3.5.

The large reorganizations, announced by GM on November 21, 2005 and by Ford on February 19, 2006, seem to suggest that these companies want to aggressively align their North American production capacity with their current production. Further erosion of their market share can then be used to build up some spare capacity to respond to sales opportunities. At the same time, through investments in flexibility OEMs will try to operate their existing capacity much more intensively than before.

---

[66] The incentive to switch your purchase decision between time periods to chase after a temporary discount is clearly larger for more expensive vehicles.

5.2   High-value production in Canada

The automobile industry is surprisingly high tech. The five biggest companies (GM, Toyota, Ford, DaimlerChrysler and Volkswagen) combined spent a total of $33.7 billion USD on R&D in 2004! There is not a single other industry with five firms spending an average of almost $7b on R&D. One can count the number of industries with *any* firm spending that much on one hand. R&D spending by suppliers is also increasing rapidly. By 2005, the automotive firm holding most U.S. patents was Robert Bosch GmbH, now the largest OEM supplier worldwide. The Ford Motor Company was runner-up. A search of the NBER patent database reveals that these two companies combined hold more than 11,000 patents and indirectly through subsidiaries countless more.

R&D is not only concentrated by firm, it is also predominantly carried out in Michigan. Its importance is not only apparent from the location of headquarters (more than 50% of the 150 largest North American suppliers are located there), but also in terms of recorded R&D spending. The Michigan Automotive R&D directory estimates that in 1999 total R&D spending in Michigan totalled US$18b, almost all of which was privately funded, and involved 65,000 employees. Only California performs more R&D, but on a per-capita basis Michigan is unrivalled in the United States[67]. 70% of the research, US$13.1b, was on automotive applications and Michigan alone represents 85% of total U.S. R&D spending in the industry.

It is not immediately obvious how Canada will attract a piece of the research pie. The vast majority of Canadian suppliers is in favor of government support for R&D activities; see below. Recent capital investment subsidies under the (federal) Canadian Skills & Innovation Project and the Ontario Automotive Investment Strategy have tied funding to locating some innovative activities in Canada. Ford is adding a research centre to its Oak-

---

[67] At $180,000 per capita of R&D, Michigan is only rivaled by Massachusetts (#3 at $151,000) and tiny states with many headquarters like Delaware (#2) or Rhode Island (#4).

ville assembly plant. GM's Beacon project features heavy investments in human capital. Canada's largest automotive firm, Magna International operates a total of 60 R&D centres and test facilities, but only 8 in Canada. While regrettable for Canada, it is hard to imagine it being otherwise as only a quarter of its production facilities is located in Canada. In terms of policy, all the government can do is to create an environment conducive to R&D such that firms that reorganize have an incentive to locate research activities in Canada. Given the generous tax treatment of R&D in Canada, current policy seems adequate[68]. It is certainly not obvious how trade policy can play a role.

In terms of high value-added production activities, Canada is keeping up better with the U.S. in the assembly sector than in parts. Table 5.1 compares the productivity record of the two countries in both automotive sectors. In the top panel, one can see that shipments per employee are equally large in both countries, in excess of $1 million CDN. The fraction of value added is similar as well and slightly higher in 2002 than in 2000. Value added per employee was higher in Canada in 2000 and only slightly lower in 2002. The fraction of production workers is a bit lower in the U.S., which increases the 2002 U.S. labour productivity advantage slightly, but at $433,574 (U.S.) versus $405,963 (Canada) the values are extremely high. The significantly higher salary for U.S. production workers, 42% higher in 2002 (31% in 2001), is almost entirely the result of a much larger share of value added being paid out to workers in the U.S. than in Canada. The U.S. industry employs more and better paid salaried employees, but that explains only a small fraction of the gap in value added going to wages, 16.0% in Canada versus 24.5% in the U.S. The much vaunted lower wage cost in Canada, courtesy of the nationally funded health care system, seems to benefit predominantly the employer.

---

[68] See Van Biesebroeck (2006), "Impediments and Facilitators to Technology Adoption. A literature survey", report prepared for Industry Canada.

Table 5.1 Productivity comparison: Canada versus the U.S.

	Canada			US		
	2000	2002		2000	2002	
Assembly (NAICS 3361-3362)						
Shipments per employee	$1,147,471	$1,141,529	-0.5%	$1,079,286	$1,191,334	10.4%
VA as % of shipments	28.8%	30.9%		26.8%	30.7%	
Value added per employee	$330,728	$352,655	6.6%	$289,188	$365,156	26.3%
Prod workers as % of total employment	84.3%	86.9%		85.0%	84.2%	
Value added per production worker	$392,246	$405,963	3.5%	$340,391	$433,574	27.4%
Production wages as % of value added	14.1%	14.0%		21.3%	18.6%	
Production wages per production worker	$55,445	$56,658	2.2%	$72,490	$80,562	11.1%
Payroll as % of value added	17.2%	16.9%		26.3%	22.6%	
Average salary of white collar workers	$65,111	$79,303	21.8%	$95,346	$92,536	-2.9%
Parts (NAICS 3363)						
Shipments per employee	$330,460	$334,145	1.1%	$378,133	$433,291	14.6%
VA as % of shipments	40.1%	36.9%		41.3%	42.4%	
Value added per employee	$132,521	$123,186	-7.0%	$156,337	$183,686	17.5%
Prod workers as % of total employment	87.2%	85.3%		80.1%	79.5%	
Value added per production worker	$152,001	$144,496	-4.9%	$195,192	$230,983	18.3%
Production wages as % of value added	29.6%	32.7%		31.1%	27.3%	
Production wages per production worker	$44,959	$47,261	5.1%	$60,699	$63,071	3.9%
Payroll as % of value added	36.1%	40.7%		41.8%	37.2%	
Average salary of white collar workers	$67,706	$66,829	-1.3%	$84,088	$89,204	6.1%

Note: all figures in CDN dollar

Source: Own calculations based on DesRosiers Automotive Yearbook (2005) and data from Statistics Canada and U.S. Census Bureau

The situation is notably different in the parts sector. The salary gap between U.S. and Canadian production workers is similar 33.5% higher in the U.S. in 2002 (35% in 2001), but here it is mainly driven by much less value added generated in Canadian firms. The share of value added paid out in wages and benefits is comparable across the two countries. The fact that Canadian firms have much fewer salaried employees makes the difference in labour productivity measured as value added per production worker –particularly stark; it was 60% higher in the U.S. in 2002.

A couple of caveats are required to put this comparison in perspective. First, a larger fraction of the parts sector output stays in Canada and the very low Canadian—U.S. dollar exchange rate in 2002 undervalues Canadian output in that year. Second, the mix within the parts sector is disadvantageous for Canada. A greater fraction of U.S. employment is in engine production which is highly capital intensive, which biases U.S. value added upward[69]. In addition, the Canadian engine sector was operating in 2002 at approximately 50% of its usual value added per worker. Third, the Canadian industry is reallocating its parts employment towards a number of sub-sectors with higher than average value added per worker: engine & engine parts and interiors, while maintaining a large employment share in a third high value added sector—transmissions (see statistics in Table 4.5).

Finally, in Table 5.2 we present a breakdown of total value added generated in the U.S. automotive industry in different years, as estimated and predicted by the Center for Automotive Research in Michigan[70]. The most important sub-sector throughout was parts and components. Increased cost pressure, due to import competition and purchasing plans of OEMs, depressed its share slightly in 2000 to 56.0%. Increased use of electronics is predicted to raise its share to 60.1% of total value

---

[69] Given that the two countries use a different breakdown of the NAICS industry classification below the "Parts and accessories manufacturing" (NAICS 3363) level, it is impossible to control for the mix of industries to make the value added comparison.

[70] Center for Automotive Research (2002), "Estimating the New Automotive Value Chain," a study prepared for Accenture.

added generated in the industry by 2010. The share of value added generated by OEMs (including assembly) is predicted to decline considerably from its 2000 high to 26.4%, slightly more than one quarter of the industry. For these firms, assembly wages are predicted to decline from 35% of their value added in 1990 to 23% in 2010. Clearly, high value added activities in the automotive sector increasingly means parts and R&D.

Table 5.2 Breakdown of the total value added generated by automotive sub-sector

(billion USD)	1990	2000	2010
Total automotive sector (U.S.)	**$291.0**	**$432.0**	**639.5**
Distribution	$36.0	$43.0	$64.6
(Advertising – dealers – freight)	(12.4%)	(10.0%)	(10.1%)
Vehicle manufacturers – wages	$25.2	$31.2	$39.1
	(8.7%)	(7.2%)	(6.1%)
Vehicle manufacturers — other value added (design, R&D, investment)	$46.2	$97.2	$129.9
	(15.9%)	(22.5%)	(20.3%)
Parts & components	$169.7	$241.9	$384.2
	(58.3%)	(56.0%)	(60.1%)
Other material inputs	$13.9	$18.7	$29.5
(energy, warranty,…)	(8.2%)	(4.3%)	(4.6%)

5.3   Policy

To gauge the importance of different policy options for the industry, it is useful to take a look at the answers Canadian parts suppliers gave to the previously mentioned APMA survey. On a scale from 1 to 7 firms were asked to rate the usefulness of different policy initiatives on a list of 20. The results of this survey are in Table 5.3. The first column indicates the number of suppliers that find the initiative useful (more than moderately so); the second column is the fraction of respondents that find the policy initiative "very useful" or rate its usefulness "extremely high"; and the third column sums the two groups. The different initiatives are organized in order of total support — any answer from 5 to 7.

335

Table 5.3 How useful to you find the following policy initiatives?

"Industry observers have suggested policies that the government could take to facilitate or enhance the growth of the Canadian auto industry. Several of those proposals are listed below. From the perspective of your firm, please rate the usefulness of these government policy proposals from 1 to 7 based on the following scale:"

	POLICY INITIATIVES:	Fraction of positive ratings		
		5	6-7	5-7
1.	Increase funding and/or tax incentives for R&D and innovation	7.1	78.6	85.7
2.	Increase incentives to domestic investors	14.3	71.4	85.7
3.	Expedite transportation infrastructure upgrades	7.1	71.4	78.5
4.	Increase incentives to foreign investors	21.4	57.1	78.6
5.	Increase funding for Technology Partnerships type programs	28.6	42.9	71.4
6.	Implement electronic border clearing system compatible with US Customs	14.3	50.0	64.3
7.	Remove tax and other barriers that slow domestic industry consolidation	14.3	50.0	64.3
8.	Change tax law to permit more rapid depreciation of new equipment	16.7	41.7	58.3
9.	Assistance for implementing productivity-enhancing equipment/systems	33.3	25.0	58.3
10.	Renew emphasis on government-industry partnerships and task forces	21.4	35.7	57.1
11.	Increase funding for auto-sector related technical education	21.4	35.7	57.1
12.	Provision of capital to facilitate new international joint ventures	14.3	42.5	56.8
13.	Increase incentives to firms using alternative energy	21.4	21.4	42.9
14.	Government-led marketing/branding initiatives focusing on the auto sector	28.6	14.3	42.9
15.	Make the use of anti-dumping/countervail legislation easier	16.7	25.0	41.7
16.	Facilitation of Canadian auto sector in rapidly growing markets	8.3	25.0	33.3
17.	Increase tax credits for firms that implement retraining/'reskilling' programs	0.0	21.4	21.4
18.	Reduce immigration restrictions on young, technologically skilled workers	0.0	21.4	21.4
19.	Rescind Canadian ratification of the Kyoto Protocol	0.0	15.4	15.4

Notes: scale of usefulness: 1 = extremely low, 4 = moderate, 7 = extremely high

84 respondents

Source: "The East Asian Automobile Industry: Opportunity or Threat? Results of a Survey of the Canadian Auto Parts Manufacturers," Canada in Asia report, January 2005

It is clearly noticeable that R&D support and investment incentives figure high on the list. These issues get almost universal approval—86% of firms support such initiatives and almost 80% of firms strongly support R&D support. On the other hand, a number of issues that have received prominent attention in the media, do not carry much industry support. For example, Canada's ratification of the Kyoto emission abatement agreement is not perceived as much of a problem. Providing support for tech-

nical education carries some support, 57% of firms, but reducing immigration restrictions on young skilled workers is not seen as a solution to the industry's shortage of skilled trades people.

In Figure 5-1, we group the average support of different initiatives by policy area. For example, investment incentives for domestic firms and FDI are grouped together. Support for R&D and funding for Technology Partnership type programs are grouped under innovation policies. The pattern that appears is quite striking. Given that only a single policy would be assigned a different priority when we use strong support instead of total support, we will use total support numbers in the discussion[71].

Figure 5-1 Ranking of interventions by policy area

What government interventions do Canadian suppliers want?
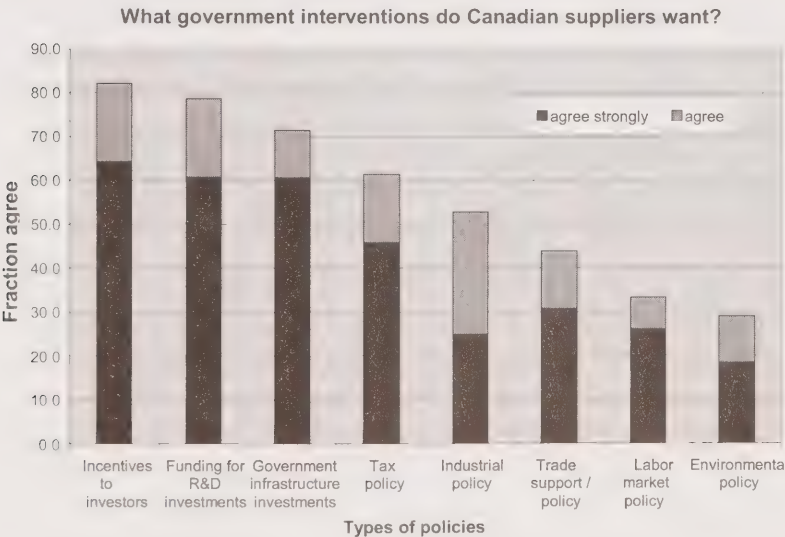
Investment support of different sorts is by far the preferred form of government intervention. More than 80% of firms think it would be a good idea for Canada to provide direct investment incentives; the level of support is only slightly higher for incen-

---

[71] While 52.8% of firms support some form of industrial policy, only 25.0% finds these very useful. Limit attention to firms that strongly support a policy, industrial policy would become second least importance.

tives for domestic than for foreign firms. In addition, assistance for firms that innovate is also widely supported. This can take the form of direct funding support, tax incentives, or funding for Partnership Programs. Finally, government investment in infrastructure, electronic border clearing with the U.S. or transportation infrastructure, is also supported by 71.4% of firms.

The next policy area receives a full 10% less support and even 15% less if we only count strong supporters. Even more striking is that the next priority is tax policy. An area of government policy that is likely to differ only in implementation from the more pro-active forms of investment support in the first three areas. The remaining policy areas, industrial, trade, labor market, or environmental policy all carry much lower levels of support. The only popular trade policy (Provision of capital to facilitate new international joint ventures) is again a form of investment support. Restricting import competition by facilitating the use of antidumping measures as well as more active government support to facilitate Canadian exports in rapidly developing markets do not gather much enthusiasm.

The overall picture that emerges is that the only interventions that carry widespread support are government investments or support for private investments. With respect to trade policy, these sentiments from the industry are reinforced by much of the analysis in the preceding chapters. Current tariff levels are sufficiently low that they are not viewed as very important policy tools. In the analysis, consumer gains from lowering tariff levels counteract producer losses with minimal net effect on Canada as a whole. In addition, the dominance of the U.S. as trading partner for the industry further reduces any effect of trade policy.

5.4   Conclusion

The sentiment towards government policy in this industry is not entirely at odds with the bulk of the results we have uncovered in this report. To summarize:

- One of the greatest changes in the industry, the changing vertical organization, is almost entirely beyond government influence. Increasing technological intensity is driven by consumer demand. Proliferation of vehicles has lead firms to adopt flexible technology and environmental awareness has spurred the development of alternative fuel vehicles. While firms clearly benefit if the government covers part of their R&D costs, the greatest success story to date, the Toyota Prius, hardly benefited from government subsidies. It remains highly doubtful to what extent government intervention can successfully steer the industry. (Section 1)

- Decreasing or abolishing import tariffs on final vehicles will benefit consumers (somewhat) and hurt Canadian production (somewhat). The net effect on welfare is likely to be very small and actual estimates, like the ones we presented in this report, will be sensitive to modeling assumptions. (Section 2)

- The only candidate firm for near term investments in new assembly capacity in Canada is Nissan. Trade policy is likely to be of limited impact in securing such investment. Infrastructure or direct investment support are likely to be much more important. The Ontario government has already started talks. (Section 3)

- The parts sector is much more vulnerable to exchange rate fluctuations, raw material prices, and bankruptcies of large firms, factors largely beyond the Canadian government's control. The very limited export success of the Canadian industry beyond the U.S. is unlikely to be to a large extent the result of trade restrictions[72]. Of course, at the margin every-

---

[72] Japan, the largest market after the U.S. and distinctly high-cost, does not charge any import tariffs, but Canadian producers have not made significant inroads.

thing helps. Current trade protection for the parts sector in Canada is very low. If giving up the limited protection that exists would lower overseas trade barriers (which tend to be higher), the net effect is likely to be positive. (Section 4)

- When asked about preferred government interventions, a large majority of firms in this industry refer to investment, R&D, and infrastructure support. (Section 5)